

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2003, 11:19:48 ; Search time 5038 Seconds
(without alignments)
3386.128 Million cell updates/sec

Title: US-09-469-200D-2
Perfect score: 2145
Sequence: 1 MRLKNIITVAFSIFWLL.....KLYSLFTIRNADWGTREKLL 417

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09469200/runat_02102003_090348_22777/app_query.fasta.1.583
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09469200.ecgn_1_13508.erunat_02102003_090348_22777 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6
-FGAPOP=7 -YGAPOP=10 -YGAPOP=0.5 -DELOP=6 -DELEXT=7

Database :
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_cm.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_nam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_hgo_hum.*
40: em_hgo_mus.*
41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2145	100.0	1254	1	AF023876	Streptococ
2	2145	100.0	1254	6	BD087249	Streptococ
3	2145	100.0	3552	1	AY173078	Streptococ
4	2138	99.7	7268	1	AF347022	Streptococ
5	1695.5	79.0	3466	1	SUB242946	Streptococ
6	1588.5	74.1	1350	1	STRHASAO	Streptococ
7	1588.5	74.1	10851	1	AE010123	Streptococ
8	1586.5	74.0	1440	1	STRHASA	Streptococ
9	1586.5	74.0	1512	6	AR232201	Streptococ
10	1586.5	74.0	9844	1	AE006637	Streptococ
11	1585.5	73.9	53895	1	AE014172	Streptococ
12	1585.5	73.9	323825	1	AP005146	Streptococ
13	1052.5	49.1	181654	1	AF065404	Bacillus
14	1052.5	49.1	181677	1	AE011190	Bacillus
15	745	34.7	1893	12	AF518732	Synthetic
16	502	23.4	6544	1	AF284858	Bradyrhiz
17	493.5	23.0	110000	1	REU80928_0	Rhizobium e
18	484.5	22.6	10973	1	AF522456	Sinorhizo
19	479	22.3	346820	1	AP003008	Mesorhizo
20	477	22.2	16124	1	RSU53327	Rhizobium s
21	476.5	22.2	6014	1	RGND	R-galegae n
22	476	22.2	152050	1	ML0672113	Mesorhizo
23	475	22.1	3373	1	RNODABC	Rhizobium m
24	475	22.1	3403	1	RTNODABCD	R-tropici n
25	474	22.1	3304	1	AF038577	Sinorhizo
26	474	22.1	4230	1	SYLNOD	Sym megapla
27	474	22.1	14188	1	AE007237	Sinorhizo
28	472	22.0	33038	1	AF222753	Bradyrhiz
29	464	21.6	1388	1	RLNODC	Rhizobium l
30	464	21.6	3094	1	RHNMNODABC	R-fredil no
31	463	21.6	10062	1	AE000076	Rhizobium
32	463	21.6	320040	6	A79350	Sequence 1
33	463	21.6	320040	6	A93002	Sequence 1
34	460	21.4	3080	1	RSNOD2	Rhizobium s
35	454.5	21.2	2627	1	RLNODABC	Rhizobium l
36	454.5	21.2	12055	1	RLNOD	Rhizobium l
37	451	21.0	4317	1	ACNODI	A-caulidans
38	451	21.0	9549	1	RHANODABC	Azorhizobiu
39	447.5	20.9	1994	9	AF234839	Homo sapi
40	447.5	20.9	4049	9	AF232772	Homo sapi
41	442	20.6	2465	10	AB097569	Rattus no
42	441.5	20.6	2370	5	AF190742	Danio rer
43	440	20.5	1659	4	AB055979	Oryctolag
44	440	20.5	1665	6	E30973	Hyaluronic
45	440	20.5	5919	10	MMU86408	Mus musculu

ALIGNMENTS

RESULT 1

```

AF023876      1254 bp      DNA      linear      BCT 15-OCT-2002
LOCUS      Streptococcus equisimilis hyaluronan synthase gene, complete cds.
DEFINITION
ACCESSION      AF023876
VERSION      AF023876.1 GI:2655099
KEYWORDS
SOURCE      Streptococcus dysgalactiae subsp. equisimilis
ORGANISM      Streptococcus dysgalactiae subsp. equisimilis
Bacteria: Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
1 (bases 1 to 1254)
Kumari, K. and Weigel, P. H.
Molecular cloning, expression, and characterization of the
authentic hyaluronan synthase from group C Streptococcus
equisimilis
J. Biol. Chem. 272 (51), 32539-32546 (1997)
JOURNAL
MEDLINE      98070433
PubMed      9405467
REFERENCE      2 (bases 1 to 1254)
AUTHORS      Weigel, P. H. and Kumari, K.
TITLE      Direct Submission
JOURNAL      Submitted (10-SEP-1997) Biochemistry & Molecular Biology.
University of Oklahoma Health Science Center, P.O. Box 26901, BMSB
860, Oklahoma City, OK 73190, USA
LOCATION/Qualifiers
1. .1254
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="genomic DNA"
/sub_species="equisimilis"
/db_xref="taxon:119602"
1. .1254
/codon_start=1
/transl_table=1
/product="hyaluronan synthase"
/protein_id="AAB8784.1"
/db_xref="GI:2655100"
/translation="MKILKILITVVAISFWLLIYVNVYLFSAKGSIIYGFLLIAY
LLVMSLFFPKFGRAGYKVAALIPSYNEDAESLETLKSVQOQTYPLAEIYVD
DGADSETKRLEDYRDGTDLSSNVIHRSEKQGRHAQWAFERSDADVFYTVDS
DTVTPDALLEELKTFNDPTFAATGHNVRNROTLLRLTDYDNFVGRSAQS
VTGNILVSGPVSIVREVVNIDRYINOTFLGIVSIGDRCRLNTATDLGKTVIQ
STAKCTDVPDKMSYLYKQNRNKSFFRESISVKKIMNPFVALWILVSNFMKL
VYSVDFVYVREDRLVLAELVLIYFALCKRIHMLKPLSLSPFYGVILHLF
VLOPLFVSLFTIRNADWGRKKLL"
BASE COUNT      335 a      240 c      261 g      418 t
ORIGIN

Alignment Scores:
Pred. No.:      8,47e-183      Length:      1254
Score:      2145.00      Matches:      417
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      100.00%      Indels:      0
DB:      1      Gaps:      0

US-09-469-200d-2 (1-417) x AF023876 (1-1254)

QY      1 MetArgThrLeuLysAsnLeuLeThrValValAlaPheSerIlePheTrpValLeu 20
Db      1 ATGAGAACATTAATAAAACCTCAACTGTTGTGGCTTAGTATTTTGGTACTGTTG 60
QY      21 IleTyrValAsnValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeu 40
Db      61 ATTACGTCATATGTTATCTCTTTGGGTAAAGGAAGCTTGCAATTTATGCGCTTTTG 120
QY      41 LeuIleAlaTyrLeuLeuValLysMetSerLeuSerPhePheTyrLysProPheLysGly 60
Db      121 CTGATAGCTTACCTATTATGTCAAATGTCCTTATCTCTTTTACAGCCATTAAGGGA 180
QY      61 ArgAlaGlyGlnTyrLysValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSer 80
Db      181 AGGGGTGGGCAATATAAGGTGTCAGCCATTATTCCTCTTATACGAAGATGCTGAGTCA 240

```

RESULT 2
 BD087249
 LOCUS
 DEFINITION
 Hyaluronan synthase gene and utilization thereof.
 ACCESSION BD087249

1254 bp DNA linear PAT 27-AUG-2002
 BD087249
 Hyaluronan synthase gene and utilization thereof.
 BD087249

81 LeuLeuGluThrLeuLysSerValGlnGlnThrTyrProLeuAlaGluIleTyrVal 100
 241 TTCTAGAGACCTTAAAGAGTGTTCAGCAGCAACCTATCCCTAGCAGAAATTTATGTT 300
 101 ValAspAspGlySerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgAsp 120
 301 GTTGACGATGAAGTGTGATGAGACAGGTATTAAAGCGCATTAAGACTATGCGGTGAC 360
 121 ThrGlyAspLeuSerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArg 140
 361 ACTGGTGAACCTATCAAGCAATGTCTATTCATCGGTCAAGAAAAATCAAGAAAGCGT 420
 141 HisAlaGlnAlaTyrPalaPheGluArgSerAspAlaAspValPheLeuThrValAspSer 160
 421 CATGCACAGGCTGGCGCTTTGAAAGATCAGACGCTGATGCTTTTTCGCGGTGACTCA 480
 161 AspThrTyrIleTyrProAspAlaLeuGluGluLeuLeuLysThrPheAsnAspProThr 180
 481 GATACCTATATCTACCTGATGCTTTAGAGAGTGTGTTAAAAACCTTTAATGACCAACT 540
 181 ValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuThrArg 200
 541 GTTTTGTGCGACGGGTCACTTAAATGTCAGAAATAGACAAACCAATCTCTTAACAGC 600
 201 LeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThr 220
 601 TTGACAGATATTCGCTATGATATGCTTTGGCGTTGACGAGCTGCCCAATCGTTACA 660
 221 GlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgGluValValPro 240
 661 GGTAAATATCCTTGTGCTCAGGTCGGCTTAGCGTTTACAGACGCGAGGTGTGTCCT 720
 241 AsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAsp 260
 721 AACATAGATAGATACATCAACACGACCTTCTCGGTATCTCTGAAGTATTTGGTATGAC 780
 261 ArgCysLeuThrAsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLys 280
 781 AGGTGCTTGACCAACTATGCACTGATTTAGGAAGACTGTTTATCAATCACTGCTAAA 840
 281 CysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTyrAsn 300
 841 TGTATTACAGATGTTCTCGAAGATGCTACTTACTTGAAGCAGCAAAACCGCTGGAAC 900
 301 LysSerPhePheArgGluSerIleLeuSerValLysLysIleMetAsnAsnProPheVal 320
 901 AAGTCTCTTTAGAGAGTCCATATTTCTGTTAAGAAATCATGACAACTCTTTGTA 960
 321 AlaLeuThrTrpThrIleLeuGluValSerMetPheMetMetLeuValTyrSerValAsp 340
 961 GCCCTATGGACCATACTTGAGGTGCTCTATGTTTATGATGCTTTTATTTCTGTGGTGGAT 1020
 341 PhePheValGlyAsnValArgGluPheAspTrpLeuArgValLeuAlaPheLeuValIle 360
 1021 TTCTTTGAGGCAATGTCAGAGAAATTTGATGGCTCAGGCTTTAGCCTTTCTGTTGAT 1080
 361 IlePheIleValAlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPhe 380
 1081 ATCTTCATGTTGTCCTGTGTCGAACATTCATTACATGCTTAAGCACCCTGCTCTTC 1140
 381 LeuLeuSerProPheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyr 400
 1141 TTGTTATCTCCGGTTTATGGGGTGTGTCATTTGTTGTCTACAGCCCTTGAATATAT 1200
 401 SerLeuPheThrIleArgAsnAlaAspTrpGlyThrArgLysLysLeuLeu 417
 1201 TCTCTTTTACTATTAGAAATGCTGACTGGGGAACACGTAATAAAATATTATTA 1251

```

VERSION      BD087249.1 GI:22632859
KEYWORDS
SOURCE       JP 2001521741-A/1.
ORGANISM     Streptococcus dysgalactiae subsp. equisimilis
              Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
              Streptococcus.
REFERENCE    1 (bases 1 to 1254)
AUTHORS     Weigel,P.H., Kumari,K. and Deangelis,P.
TITLE       Hyaluronan synthase gene and utilization thereof
JOURNAL     Patent: JP 2001521741-A 1 13-NOV-2001;
              THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA
COMMENT     OS Streptococcus equisimilis
              PN JP 2001521741-A/1
              PD 13-NOV-2001
              PF 30-OCT-1998 JP 2000519083
              PR 31-OCT-1997 US 60/064435, 26-OCT-1998 US 09/178851 P1
              PC PAUL H WEIGEL, KSHAMA KUMARI, PAUL DEANGELIS
              C12N15/09, A61K31/728, A61P43/00, C12N1/21, C12P19/04, PC
              C12Q1/68//
              PC (C12N15/09, C12R1:46), C12N15/00, (C12R1:46) CC
              Hyaluronan synthase gene and utilization thereof FH Key
FEATURES     Location/Qualifiers
FT           source 1..1254
              /organism="Streptococcus equisimilis".
              Location/Qualifiers
              1..1254
              /organism="Streptococcus dysgalactiae subsp. equisimilis"
              /mol_type="genomic DNA"
              /sub_species="subsp. equisimilis"
              /db_xref="taxon:119602"
BASE COUNT   335 a 240 c 261 g 418 t
ORIGIN
Alignment Scores:
Pred. No.:   8.47e-183 Length: 1254
Score:       2145.00 Matches: 417
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB:          6 Gaps: 0

US-09-469-200D-2 (1-417) x BD087249 (1-1254)

QY 1 MetAtrThrLeuLysAsnLeuIleThrValValAlaPheSerIlePheTrpValLeuLeu 20
Db 1 ATGAGAACATTAAACCTCATACACTGCTGGCCCTTAGTATTATTTTGGTACGTG 60
QY 21 IleTyrValAsnValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeu 40
Db 61 ATTACGTCATATGTTTATCTCTTGGTGCTAAAGGAAGCTTGTCAATTTATGGCTTTTG 120
QY 41 LeuIleAlaTyrLeuLeuValLysMetSerLeuSerPheTyrLysPropheLysGly 60
Db 121 CTGATAGCTTACCTATTAGTCAAAATGCTTATCTCTTTTACAGCCATTATAGGA 180
QY 61 ArgAlaGlyGlnTyrLysValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSer 80
Db 181 AGGCTGGCAATTAAGTTGACGCAATATTCCCTCTATATAAGGAAGATGCTGAGTCA 240
QY 81 LeuLeuGluThrLeuLysSerValGlnGlnThrTyrProLeuAlaGluIleTyrVal 100
Db 241 TTGCTAGAGACCTTAAAGAGTTTCAGCAGCAACCTATCCCTAGCAGAAATTTATGTT 300
QY 101 ValAspGlySerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgasp 120
Db 301 GTTGACATGGAGGTGCTGATGAGACAGAGTATTAGCCGATTTAGACATATGCGTGAC 360
QY 121 ThrGlyAspLeuSerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArg 140
Db 361 ACTGGTGACCTATCAAGCAATGTCAATTTTCATCGGTTCAGAGAAAATCAAGGAAGCGT 420
QY 141 HisAlaGlnAlaTrpAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSer 160
              161 AspThrTyrIleTyrProAspAlaLeuGluGluLeuLeuLysThrPheAsnAspProThr 180
              181 ValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArg 200
              541 GTTTTGTGCGCAGCGGTCCCTTAATGTCAGAAATAGACAAACCAATCTCTTACACGC 600
              201 LeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThr 220
              601 TTGACAGATATTCGCTATGATGCTTTTGGCGTTGAACGAGCTGCCCAATCCGTTACA 660
              221 GlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgGluValValPro 240
              661 GGTAATATCTTGTTCGCTCAGGTCCGCTTAGCGTTTACAGCGGAGGTGGTTGTCCT 720
              241 AsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAsp 260
              721 AACATAGATAGATACATCAACAGACCTTCCTGGGTATTCCTGTAAGTATGGTGATGAC 780
              261 ArgCysLeuThrAsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLys 280
              781 AGTGTCTTGACCAACTATGCAACTGATTTAGGAAGAGACTGTTTATCAATCCACTCTAAA 840
              281 CysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrpAsn 300
              841 TGTATTACAGATGTTCTTCGACAGATGCTCTACTTACTTGAACGCGAANACCGTGGAC 900
              301 LysSerPhePheArgGluSerIleIleSerValLysLysIleMetAsnAsnPropheVal 320
              901 AAGTCCCTCTTTAGAGAGTCCATATTCTGTTAAGAAATCATGAACAATCCCTTTGTA 960
              321 AlaLeuThrThrIleLeuGluValSerMetPheMetMetLeuValTyrSerValValAsp 340
              961 CCCCTATGGACCATCTCTGAGGTGCTATGTTTATGATGCTGTTTATTTCTGGTGGAT 1020
              341 PhePheValGlyAsnValArgGluPheAspTrpLeuArgValLeuAlaPheLeuValIle 360
              1021 TTCCTTTAGGCANTGTCAGAGATTTGATGGCTCAGGGTTTACCTTTCTGGTGATT 1080
              361 IlePheIleValAlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPhe 380
              1081 ATCTTCATCTTGCCTCTGTCGGAACATTCATTCATGCTTAAGCACCCGCTGCTCTTC 1140
              381 LeuLeuSerProPheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyr 400
              1141 TTGTTATCTCCGTTTATGGGTGCTGCAATTTGTTTGTCTACAGCCCTTGAATATATAT 1200
              401 SerLeuPheThrIleArgAsnAlaAspTrpGlyThrArgLysLysLeuLeu 417
              1201 TCTCTTTTACTATTAGAAATGCTGACTGGGACACGTAATAAAATATATTA 1251

RESULT 3
AY173078 3552 bp DNA linear BCT 15-JAN-2003
LOCUS Streptococcus equi subsp. zooepidemicus hyaluronic acid operon,
DEFINITION partial sequence.
ACCESSION AY173078
VERSION AY173078.1 GI:27762599
KEYWORDS
SOURCE Streptococcus equi subsp. zooepidemicus
ORGANISM Streptococcus equi subsp. zooepidemicus
          Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
          Streptococcus.
REFERENCE 1 (bases 1 to 3552)
AUTHORS Wu, X. and Gao, H.
TITLE Molecular characterization of hyaluronic acid operon from group C
JOURNAL Streptococcus equi subsp. zooepidemicus
          Unpublished
REFERENCE 2 (bases 1 to 3552)
          Wu, X., Gao, H. and Chen, J.

```

TITLE Molecular cloning, expression, and characterization of hasC and galP, two genes encoding UDP-glucose pyrophosphorylase in group C Streptococcus equi subsp. zooepidemicus

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 3552)

AUTHORS Wu, X. and Gao, H.

TITLE Direct Submission

JOURNAL Submitted (05-NOV-2002) Department of Biology, Tsinghua University, Beijing 100084, China

FEATURES

source

1..3552 Location/Qualifiers

/organism="Streptococcus equi subsp. zooepidemicus"

/mol_type="genomic DNA"

/sub_species="zooepidemicus"

/db_xref="taxon:40041"

1..3552

/gene="hyaluronic acid operon"

1..1254

/gene="hasA"

1..1254

/gene="hasA"

/note="required for hyaluronic acid synthesis"

/codon_start=1

/transl_table=11

/product="hyaluronan synthase"

/protein_id="AAO20097.1"

/db_xref="GI:27762600"

/translation="MTLKNLTVVAFSIFWVLLIYVNVYLFEGKSGLSIYGLLIAY

LVKMSLEFLKPKRAGQYKVAALIFSYNEDAESLLETKSVQCCQYPLAELIYVD

DGSADETKKRIEDYVRDQGLSSNVIVHRSEKNGKRAQAWAFERSDADVLTVD

DTYIDALELLKTFNDPTFAATGHLNVRNQTNLRLTDIRDNAGFVERAAQS

VTGILVSGSLVYREVYVNIIDRYNFTFLGIPVSGIDRCLTNYATDLGKTYIQ

STAKCTIDPKMGTYLKKQCNWKSFFRESISVKIMNPNFVALWTILEVSNFNL

VYSVDFFVGNVREDLRLVLAFLVLIIVLALCRNIHYMKLHPLSLSPYGVHLF

VLOPLKLSLETIRNADGTRKLL"

1507..2712

/gene="hasB"

1507..2712

/gene="hasB"

/codon_start=1

/transl_table=11

/product="UDP-glucose dehydrogenase"

/protein_id="AAO20098.1"

/db_xref="GI:27762601"

/translation="MKISVAGSYVGLSILLAQHNDVTVVDIIDEKVRLLINGLSP

IKDADTEYLKNAPLNLATLDGASAYSNADLIILATPTNIDSEKNYEDTRHVEVIE

QVLDLNASATIIILKSTIPLGFIKHREKYQDRLIFSPFELRESKALYDNLVPSRLIV

SYEKDSRVIOAKAFAGLLKEGAKSDTPVLPWGSQEAELVKNFLANRYSYF

NELDTYSKGLDQQRVEGVCHDQRIGNHNNPSPFGYCLPKDQKLLANRYGIP

QSLMSAIVESNKRKSLAQLDORASSQKQGVPLTIGFYRLIMKSNDRFRESAII

DIIDINDGVNVIYEFMLGEDIYRVYVKDLEQFKNESTIIVSNRFBDDLDGVIDKV

YTRDVFGRD"

2762..3552

/gene="hasC"

2762..3552

/gene="hasC"

/codon_start=1

/transl_table=11

/product="UDP-glucose pyrophosphorylase"

/protein_id="AAO20099.1"

/db_xref="GI:27762602"

/translation="MTKVRKAIIPAAAGLRTPLPATKALAKEMLPVOKPTIQIVEE

ALKAGTEILVYTGAKRSIEDHDFNFEYLNQAKGTELLKLVDETTAINLHFTIR

QSHPRGLGDAVLQAKAFVGNPEFVLMGDDLDITNPISAKPLTKQIEDYDCTHASII

AVMVRPHEEVSNGVIAPOGKAVKGLYSVTEVEKPSDEAPSDLAIGRVLLTPEIF

ALLENQAPGAGNEVOLITDAIDKLNKTQRYFAREFKRGYDVGKFG"

BASE COUNT 1006 a 732 c 817 g 997 t

ORIGIN

Alignment Scores:

pred. No.: 3,8e-182 Length: 3552

Score: 2145.00 Matches: 417

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 1 Gaps: 0

US-09-469-200d-2 (1-417) x AY173078 (1-3552)

QY 1 MetAgtgThrLeuLysAsnLeuIleThrValValAlaPheSerIlePheTrpValLeuLeu 20

DB 1 ATGAGAACATTAAACCACTCATAACTCTGTGTGGCCTTTAGTATATTTTGGGTACTGTG 60

QY 21 IleTyrValAsnValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeu 40

DB 61 ATTAGCTCAATGTTATCTCTTTTGGTCTAAAGAAAGCTTGTCAATTTATGCTTTTG 120

QY 41 LeuIleAlaTyrLeuLeuValLysMetSerLeuSerPhePheTyrLysPropheLysGly 60

DB 121 CTGATAGCTTACCTTATAGTCAAAATGCTTATCTCTTTTTCACAGCCATTTAAGGGA 180

QY 61 ArgAlaGlyGlnTyrLysValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSer 80

DB 181 AGGCTGGGCAATATAGGTTGCACCCATTTATCCCTCTTATACGAGATGCTGAGTCA 240

QY 81 LeuLeuGlnThrLeuLysSerValGlnGlnThrTyrProLeuAlaGluIleTyrVal 100

DB 241 TTGCTAGAGACCTTAAAGAGTGTTCAGCAGCAACCTATCCCTAGCAGAAATTTATGTT 300

QY 101 ValAspAspGlySerAlaAspGlnThrGlyIleLysArgIleGluAspTyrValArgSp 120

DB 301 GTTGACGATGGAAGTGTGTAGACAGGATTAAGCGCATTAAGCACTATGTGCGGTGAC 360

QY 121 ThrGlyAspLeuSerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArg 140

DB 361 ACTGGTGACCTATCAACCAATGTCTTCTCATCGGTCAAGAAAAATCAAGAAAGCGT 420

QY 141 HisAlaGlnAlaTyrAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSer 160

DB 421 CATGCACAGGGCTGGGCTTTTGAAGATCAAGCGCTGATGTCTTTTGACCGTTGACTCA 480

QY 161 AspThrTyrIleTyrProAspAlaLeuGluLeuLysThrPheAsnAspProThr 180

DB 481 GATCTTATATCTACCTGATGCTTTAGAGAGTGTGTTAAACCTTTATGACCAACT 540

QY 181 ValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArg 200

DB 541 GTTTTCTTGGCAGCGGTCACTTAATGTCAAGAAATAGACAAACCAATCTCTTAACACG 600

QY 201 LeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThr 220

DB 601 TTGACAGATATTCGCTATGATATATGCTTTTGGCGTTGACGAGCTGCCCAATCCGTACA 660

QY 221 GlyAsnIleLeuValLysSerGlyProLeuSerValTyrArgGluValValPro 240

DB 661 GGTAATATCTTGTGCTCAGGTCGCTTAGCGTTTACAGACGCGAGGTGTGTTGTCCT 720

QY 241 AsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAsp 260

DB 721 AACATAGATAGATACATCAACACGACCTTCTCTGGGTATTCCTGTAAGTATTTGGTGATG 780

QY 261 ArgCysLeuThrAsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLys 280

DB 781 AGGTGCTTGACCACTATGCAACTGATTTAGAAGAGCTGTTTATCAATCCATCCGTCAA 840

QY 281 CysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTyrAsn 300

DB 841 TGTATTACAGATGTTCTCTGACAAAGATGCTACTTACTTGAAGCAGCAAAACCGCTGGAAC 900

QY 301 LysSerPheArgGluSerIleIleSerValLysIleMetAsnAsnPropheVal 320

DB 901 AAGTCCTTCTTTAGAGATGCTTATTTCTGTTAAGAAATCATGACAAATCTCTTGTGA 960

QY 321 AlaLeuThrThrIleLeuGluValSerMetPheMetMetLeuValTyrSerValValAsp 340

DB 961 GCCCTATGGACCACTACTTGGAGGTGCTATGTTTATGATGCTTGTATTCTTGTGGTGGAT 1020


```

QY 341 PhePheValGlyAsnValArgGluPheAspTriPheuArgValLeuAlaPheLeuValIle 360
Db 1021 TTTCTTTGAGCAATGTCAGAGAAATTCATTGGCTCAGGCTTTAGCCTTCTGGTGAT 1080
QY 361 ilePheIleValAlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPhe 380
Db 1081 ATCTTCATTGTTGCCCTGTGTCGGAACATTCATTACATGCTTAAGCACCGCTGCTTC 1140
QY 381 LeuLeuSerProPheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyr 400
Db 1141 TTCTTATCTCGCTTTATGGGTGCTGCAATTTGTTGCTCTACAGCCCTTGAATATAT 1200
QY 401 SerLeuPheThrIleArgAsnAlaAspTrpGlyThrArgLysLysLeuLeu 417
Db 1201 TCTCTTTTACTATTAGAAATGCTGACGGGACACGTAATAAAATTAATA 1251

RESULT 4
AF347022 7268 bp DNA linear BCT 13-NOV-2002
LOCUS Streptococcus equi subsp. zooepidemicus has operon, partial
DEFINITION
ACCESSION AF347022
VERSION AF347022.1 GI:24940621
KEYWORDS
SOURCE Streptococcus equi subsp. zooepidemicus
ORGANISM Streptococcus equi subsp. zooepidemicus
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.
REFERENCE 1 (bases 1 to 7268)
AUTHORS Blank,L.M., Rugenholtz,P. and Nielsen,L.K.
TITLE Cloning and characterization of the hyaluronic acid synthesis (has) operon from Streptococcus equi subsp. zooepidemicus
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 7268)
AUTHORS Blank,L.M. and Nielsen,L.K.
TITLE Direct Submission
JOURNAL Submitted (09-FEB-2001) Chemical Engineering Department, The University of Queensland, Coopers Road, Building #74, Brisbane, Queensland 4072, Australia
FEATURES
source
1..7268
/organism="Streptococcus equi subsp. zooepidemicus"
/mol_type="genomic DNA"
/sub_species="zooepidemicus"
/db_xref="taxon:40041"
469..>7268
/gene="has operon"
469..474
/gene="has operon"
481..>7268
/gene="has operon"
/evidence=experimental
512..1784
/gene="has"
512..517
/gene="has"
531..1784
/gene="has"
/codon_start=1
/transl_table=11
/product="hyaluronic acid synthase"
/protein_id="AAN65248.1"
/db_xref="GI:24940622"
/translators="MSTLANLTVVAFSIFWLLIYVNVYLFAGKSLSYGLLIAY
LVKMSLFYKPKRAGQYKVAALIPSDNEDASLETLSVQQCTYPLAEITVD
DSADETGIRKIEDYVDTGLSSNVIVHRSKNQKRAQAFERSADDFVLTVDS
DTIYFDALIELLKTNDPTFAATGHLNVRNQTNLRTIDRYDNAGFVERAQS
VTGNLVCSGPLSVYREVVVNPIDRGINOTFLGIPVSGIDRCLNTYATDLGKTVYQ
STAKCITDVPKMSLYLKQONKNSPFRSIIISVKKIMNPFVALATILEYSMMML
VYSVDFVDNREFDMLVLAFLVLIETVALCRNIHMKPLSFLISFFYGLVILF
VLOPLKYSLPTIRNADWGTAKLL"
2023..3242
gene
/gene="hasp"
2023..2028
/gene="hasb"
2037..3242
/gene="hasb"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="UDP-glucose dehydrogenase"
/protein_id="AAN65249.1"
/db_xref="GI:24940623"
/translators="MKISVAGSGYGLSILLQAHDVTVVDIIDEKVRILNQGISP
IKDADTEYLKNAPLNLTATLDCASAYSNADLIITPTNYDSERNVPTDRHVEVIE
QVLDLNASATIIKSTIPLGFKHVKRYQDRIISPELRESKALYDNLVPSRLIV
SEYEDSPVTOAKAFAGALLKEGAKSDTVPFMGSOAEAYKLFANTFLANRYSYF
NEUDTYSKGLDQARVIEGVCHDQRIHNNFSGIGYCLPKDQKLLANRSGIP
QSLMSAIVGSNKRIRKSYLAQIILDRASSQKQGVPLTIGYRILKMSNDSNFRSAIK
DIIDIINDYGVNIVIVPEMLGIDIGYRVVKDLEQFKNESTIIVSNRPFDDIGDVIDKY
YTRDVFGRD"
3280..4194
/gene="hasc"
3280..3285
/gene="hasc"
3292..4194
/gene="hasc"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="UDP-glucose pyrophosphorylase"
/protein_id="AAN65250.1"
/db_xref="GI:24940624"
/translators="MTKVKKAIIPAGLGTFLPATKALAKEMLPVDKPTIOFIVEE
ALKAGEETLVVTCGAKRSIEDHFSNFELEYNLOAKGTELLKLVDETALNHSTR
QSHPRGLDGYLQAKAFYGVNEPFVYMLGDDLDITNSAKPLAKLIIEDICTHASTI
AVMRPHEEVSNGIYAPQGVKAGLYSVETFEKPSDFASDLAIIGRTLLTPEIF
ALLENQAPAGNEVLADAIDKLNTQRYVAREFEKRGVDVGDFGFKMYSLOYALKH
POVKDDLDYIIKLSKQINKDKVK"
4349..5744
/gene="glmU"
4349..4354
/gene="glmU"
4362..5744
/gene="glmU"
/codon_start=1
/transl_table=11
/product="UDP-N-acetyl-glucosamine pyrophosphorylase"
/protein_id="AAN65251.1"
/db_xref="GI:24940625"
/translators="MKNYAILAAGKGTMRNSGLSKVLHKVGLSMLSEVLKSVSALA
PQKLTAVIGHQAEQVRAVLGDQDLTVVQEEQLGTGHAVMABEELSGLEGTLVIAGD
PLTIRGESLKALDYHIREKNVATILTNADKPFYGRIRNAAGEVNVIVQEKDANE
AEOEVEKINTGTIFDNKRLFEALKHLTDDNAQGEVYLTDTLSIFKASQEKVAYLLK
DFDESGLVNDLALAQAIEVIMQERINKQHLNGVTLQNPATYIESSEIAPDVLEA
NVILKQTRIGRSYNTSITLDSRLGEGVYVVSQSVGSLVADGVTVGFIAHIRPD
QTDCECHVIGNEVEVKGHIGANTKAGHITYLGNAEIGSEVNIAGSITVNDGQRKY
QTVIGDHAFIGSHSTLIAPVEVGENALTAAGSTIAQSVPADSVIAGRSQVYKEGAK
RLPHHPDQPK"
5905..7268
/gene="psi"
5905..5910
/gene="psi"
5919..7268
/gene="psi"
/codon_start=1
/transl_table=11
/product="glucose-6-phosphate isomerase"
/protein_id="AAN65252.1"
/db_xref="GI:24940626"
/translators="MSHTTDSYKLEQFAGQHEIDFLQGVTEADQALROGTGPGSD
FLWLELTPNVDKFEFARILKAAEKIKADSVLAVIGIGSVYLGAKAIDFLNSHFAN
FLWLELTPNVDKFEFARILKAAEKIKADSVLAVIGIGSVYLGAKAIDFLNSHFAN
LQYAKERKAPQILLNAGNSISSVLAIDLVDYQDKDFSVNVISKSTTTERALAFVFK
ELVYKQGEANKRIYATTATDKYKAVKVEADANHWETFPVDPNPGVRSFVTEAVGLL
PIAASGADITALEMEGANAARKDLSSDKISENIATQYAVVRNLIYRKGVYTEILANVEP

```

SLOVSESEWKKOLAGESEKDKGIYPTSNFSTOLHSIGFIOEGRNLFETVIRVDK
PRQNIIPMAEDLDGGLYLGKQDVFYNNKATDGLVLAHTDGVFNPFITLPQDEF
TLGTYIFELALISGLYNGVNFDFQPGVEAYKKNMFALLKPGFELGAALNARI*
BASE COUNT 2105 a 1485 c 1733 g 1945 t
ORIGIN

Alignment Scores:
Pred. No.: 4,53e-181 Length: 7268
Score: 2138.00 Matches: 416
Percent Similarity: 99.76% Conservative: 0
Best Local Similarity: 99.76% Mismatches: 1
Query Match: 99.67% Indels: 0
DB: 1 Gaps: 0

US-09-469-200D-2 (1-417) x AF347022 (1-7268)

QY 1 MetArgThrLeuLysAsnLeuIleThrValValAlaPheSerIlePheTrpValLeuLeu 20
Db 531 ATGAGAACATTTAAACACCTCATACCTGTTGGCCCTTAGTATTTTGGGTACTGTTG 590
QY 21 IleTyrValAsnValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeu 40
Db 591 ATTTACGTCATGTTATCTCTTGGTGCTAAGGAAGCTTGTCAATTTAIGCTTTTGTG 650
QY 41 LeuIleAlaTyrLeuLeuValLysMetSerLeuSerPhePheTyrLysProPheLysGly 60
Db 651 CTGATAGCTTACCTATTAGTCAAAATGTCCTTATCCTTTTTCACAGCCATTAAAGGGA 710
QY 61 ArgAlaGlyGlnTyrLysValAlaAlaIleLeuProSerTyrAsnGluAspAlaGluSer 80
Db 711 AGGCTGGGCATTAAGTGTGACCCATATTCCCTCTTATAACGAAGAIGCTGAGTCA 770
QY 81 LeuLeuGluThrLeuLysSerValGlnGlnThrTyrProLeuAlaGluIleTyrVal 100
Db 771 TTGCTAGAGACCTTAAAGAGTGTTCAGCAGCAACCTATCCCTAGCAGAAATTTATGTT 830
QY 101 ValAspAspGlySerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgAsp 120
Db 831 GTTGACGATGAAGTGTGATGAGACAGGTATTAAAGCGATTGAAGCACTATGTGCGTGAC 890
QY 121 ThrGlyAspLeuSerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArg 140
Db 891 ACTGTGACCTTACGCAATGTCTATGTTCCCGGTGAGAAATATCAAGAAAGCGT 950
QY 141 HisAlaGlnAlaThrAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSer 160
Db 951 CATGCACAGCGCTGGCGCTTCAAAAGATCAGACGCTGATGCTTTTGGACCGTGACACA 1010
QY 161 AspThrTyrIleTyrProAspAlaLeuGluGluLeuLysThrPheAsnAspProThr 180
Db 1011 GATACCTATATCTACCTGATGCTTTAGAGAGTGTGTAACCACTTTAATGACCAACT 1070
QY 181 ValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArg 200
Db 1071 GTTTTGTGCGAGGGTCACTTAAATGTCAGAAATAGACAAACCAATCTCTTAACAGCG 1130
QY 201 LeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThr 220
Db 1131 TTGACAGATATTCGTATGATATAATGCTTTGGGCTTGAACAGCGTCCCAATCCGTTACA 1190
QY 221 GlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgGluValValValPro 240
Db 1191 GGTATATCTCTGCTGCGTCCGCGCGCTTAGCGTTTACAGACGCGAGGTGGTGTGCTCT 1250
QY 241 AsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAsp 260
Db 1251 AACATAGATAGATACATCAACAGACCTTCTGGGTATTCCTGTAATATCGGTGATGAC 1310
QY 261 ArgCysLeuThrAsnTyrAlaThrAspLeuGlyIleThrValTyrGlnSerThrAlaLys 280
Db 1311 AGGTGCTTGACCAACTATGCAACTGATTAGGAAGACGCTGTTTATCAATCCACTGCTAAA 1370
QY 281 CysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrpAsn 300

Db 1371 TGTATACAGATGTTCTCTGCAAGATGCTACTTACTTGAAGCAGCAAAACCGCTGGAAC 1430
QY 301 LysSerPheArgGluSerIleIleSerValLysLysIleMetAsnAsnProPheVal 320
Db 1431 AAGTCCTTCTTTAGAGAGTCATTTCTGTAAAGAAATCATGAACAATCCTTTGTA 1490
QY 321 AlaLeuThrTrpIleLeuGluValSerMetPheMetMetLeuValTyrSerValValAsp 340
Db 1491 GCCCTATGACCATACTTGAAGGTCTCTATGTTTATGATGCTTGTATTCTGTGGTGGAT 1550
QY 341 PhePheValGlyAsnValArgGluPheAspTrpLeuArgValLeuAlaPheLeuValIle 360
Db 1551 TTCTTTGTAGACAATGTCAGAAATTTGATTGGCTCAGGGTTTGGCCTTTCGGGTGATT 1610
QY 361 IlePheIleValAlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPhe 380
Db 1611 ATCTTCATGTTGCTCTTGTGTAATATTCATATATGCTTAAGCACCCTGCTGCTTC 1670
QY 381 LeuLeuSerProPheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyr 400
Db 1671 TTGTTATCTCCGTTTATGGGGTACTGCAATTTGTTTGTCTACAGCCCTTGAATTTGAT 1730
QY 401 SerLeuPheThrIleArgAsnAlaAspTrpGlyThrArgLysLysLeuLeu 417
Db 1731 TCTCTTTTACTATTAGAAATGCTGACTGGGAACACGTAATAAATATTATA 1781

RESULT 5

SUB242946 3466 bp DNA linear BCT 06-JAN-2001
LOCUS Streptococcus uberis hasa and hasb genes for putative hyaluronan
DEFINITION synthase and UDP-glucose dehydrogenase.
ACCESSION AJ242946
VERSION AJ242946.2 GI:7532762
KEYWORDS hasa gene; hasb gene; hyaluronan synthase; UDP-glucose
dehydrogenase.
SOURCE Streptococcus uberis
ORGANISM Streptococcus uberis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1
AUTHORS Ward, P.N., Field, T.R., Ditcham, W.G., Maguin, E. and Leigh, J.A.
TITLE Identification and disruption of two discrete loci encoding
hyaluronic acid capsule biosynthesis genes hasa, hasb, and hasc in
Streptococcus uberis
JOURNAL Infect. Immun. 69 (1), 392-399 (2001)
MEDLINE 20569205
PUBMED 11119529
REFERENCE 2
AUTHORS Ward, P.N.
TITLE Direct Submission
JOURNAL Submitted (07-JUN-1999) Ward P.N., Environmental Microbiology,
Institute for Animal Health, Compton Laboratory, Compton, nr
Newbury, Berkshire, RG20 7NN, UNITED KINGDOM
REMARK Revised by [3]
REFERENCE 3 (bases 1 to 3466)
AUTHORS Ward, P.N.
TITLE Direct Submission
JOURNAL Submitted (07-Apr-2000) Ward P.N., Environmental Microbiology,
Institute for Animal Health, Compton Laboratory, Compton, nr
Newbury, Berkshire, RG20 7NN, UNITED KINGDOM
COMMENT On Apr 10, 2000 this sequence version replaced gi:5441847.
FEATURES
source
1..3466
/organism="Streptococcus uberis"
/mol_type="genomic DNA"
/strain="Q140J"
/db_xref="taxon:1349"
/country="United Kingdom"
23..1276
/gene="hasa"
23..1276
/gene="hasa"

```

/function="putative biosynthesis of hyaluronic acid"
/codon_start=1
/transl_table=11
/product="hyaluronan synthase"
/protein_id="CAB46918.2"
/db_xref="GI:7532763"
/db_xref="SPTREMBL:O9X9L5"
/translation="MEKLNLTFFWFIPLMLIIGLVNFVFGKSLTVYGIILLY
LSIKMGLSFFRYPKSGVGQKVAALIPSNEDGVGLLETKSKVQKQYPAEIPFID
DGSVDKTKIKLVEDYKLVNGFGQVIVHQPENPVKRRHAQAFERSADVFLTVDS
TYIYPDALBELLKTFNDEVAATGHLNARNRQTNLTITLDIRYDNAGFVERAAQV
TGNLVCSPGLSIYRSGVGNLEERTVSGTFLGVPVSGDDRLNTYATDLGKTVQV
TARCDTDPDKFKVFKIQKQNRWKSFEESIISVKLLATPSPVAVWTUTEVSMILV
YSIFSLIGAEQENLIKLVAFYLIIFVALCRNWHVHKHPFAFLSPFFGLHLFV
LQPLKYSFTLRNATWGRKKYSK"
1301..2506
/gene="hasB"
1301..2506
/gene="hasB"
/function="putative biosynthesis of D-glucuronic acid"
/codon_start=1
/transl_table=11
/product="putative UDP-glucose dehydrogenase"
/protein_id="CAB46919.2"
/db_xref="GI:7532764"
/db_xref="SPTREMBL:O9X9L4"
/translation="MKIIVAGSGVGLSLSVLLAKNPVTVDLIEKKVNLINQKSP
IQVDYENIKKKLQLRAVLDAQAFROADLILIAIPTYDVEKNFFDTSHVIVIE
KALANQALLVKSITPLGFKMKQKQYOTDRILFSPFERESKALKDNLVPSRIIV
SFEDDSNEVIERAKTFAQLLKSLDKQVPLFMGSAEAEVAKLFANTYLAMRVYF
NELTYAENGRVDNIIEGVCHDRIGIHYNNPSFGYGYCLPKDKTKOLLAGVDGP
OSLKAIVDSNKIRKEYTASQILQSLSDINVDPKDATIGYRLIMKSNDSNFRSAIK
DIIDHISKYQINLVLYEPMNEDPDLPIIDLSDFKAMSHIIVSNRYDLALEDYKEV
YTRDIYGVQ"
complement(2703..>3466)
/feature="ORF1"
/codon_start=3
/transl_table=11
/product="hypothetical protein"
/protein_id="CAB87142.1"
/db_xref="GI:7532765"
/db_xref="SPTREMBL:Q9L4H1"
/translation="QAFSVLYLVLYINKKIPHRLKKDDLTLDRNLNHRALGFPPMA
FOASIIAIGAITLQVTKLGTAEIAIAIAISKTDQLAMFLMNLGLALSTLQNTYG
AKAYRILEGLKDSLIICIMWSVFAILLIFORFRFFGIFISKADPAVPHLAYLYLI
NGSLYWLIALILFRSFIQGLGSGFVETLAGIMELVMRAGVAIIGLLYFGFVGASAN
PAMIGSVLILIPSSILILIKKRODPENELRNL"
BASE COUNT 1158 a 531 c 666 g 1111 t
ORIGIN

Alignment Scores:
Pred. No.: 6,91e-142 Length: 3466
Score: 1695,50 Matches: 312
Percent Similarity: 87,47% Conservative: 51
Best Local Similarity: 75,18% Mismatches: 51
Query Match: 79,04% Indels: 1
DB: 1 Gaps: 1

US-09-469-200D-2 (1-417) x SUB242946 (1-3466)

QY 1 MetArgThrLeuLysAsnLeuLeuThrValValAlaPheSerIlePheTrpValLeuLeu 20
Db 23 ATGGAAAAATATAAAATCTCATTTATGACTTTTATTTCTCTGTCGCTCAATTT 82
QY 21 IleIyrValAsnValIyrLeuPheGlyAlaLysGlySerLeuSerIleIyrGlyPheLeu 40
Db 83 ATGGGCTTAAATGTTTTTGTATTTGGAAGTAAAGGAAGTCTAACAGTGTATGGGATTAT 142
QY 41 LeuIleAlaIyrLeuLeuValLysMetSerLeuSerPhePheIyrLysProPheLysGly 60
Db 143 CTATTAACCTATTGTCGATAAAAATGGGATTATCTTTTTTTTATCGTCCCTATAAGA 202
QY 61 ArgAlaGlyGlnIyrLysValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSer 80

```

```

|||||
203 AGTGTAGGCAATATATAGGTAGCAGCTATTATCCCATCTTATATAGAGATGGTGGT 262
QY 81 LeuLeuGluThrLeuLysSerValGlnGlnThrTyrProLeuAlaGluIleIyrVal 100
Db 263 TTACTAGAACTCTAAGAGGTGTTCAAAAACAACATATCAATTCGAATAATTCGTA 322
QY 101 ValAspAspGlySerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgAsp 120
Db 323 ATTGACCATGGGTCTAGTAGATAAACAAGCATATAAAATGGTGCAGAGACTATGTGAAGTTA 382
QY 121 ThrGlyAspLeuSerSerAsnValIleValHisArgSerGluLysAsnGlnGlySerG 140
Db 383 AATGGC---TTTGAGAGCAACAGTTATCGTTCATCAGATGCTGAAATGTTGGTAAAGA 439
QY 141 HisAlaGlnAlaIleAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSer 160
Db 440 CATGCTCAGGCTTGGGCATTTTGAAGGTCTGATGCTGATGTTCTTAAACAGTGGATTC 499
QY 161 AspThrTyrIleTyrProAspAlaLeuGluLeuLeuLysThrPheAsnAspProThr 180
Db 500 GATACCTACATCTATCTCGATGCTCTTGAAGAAATTTATAAGACATTTAATGATCCAGAG 559
QY 181 ValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArg 200
Db 560 GTCTACGCTGCAACTGGTCATTTAATGCAAGAATAAGACAACTAATCTCTTAACCTAGA 619
QY 201 LeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaGlnSerValThr 220
Db 620 CTGACTGATATCGTTAGTAGTAATGCAATTTGGTGTAGAACGTGCTGCTCAGTCTGTTCAG 679
QY 221 GlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgGluValValValPro 240
Db 680 GGAATATTTGGTTGTTCCGACCTTTAGTATTATAGACGTTCCGTCGGTATTTCCA 739
QY 241 AsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAsp 260
Db 740 AATCTTGACGGTATACCTCACAACATTTCTTGGTGTCCCTGTAAGCATAGGGGATGAC 799
QY 261 ArgCysLeuThrAsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLys 280
Db 800 CGTTGTTTGACAAATATGCAACTGATTTGGGAAAAAGCGTTTATCAGTCAACTGCAAGA 859
QY 281 CysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrpAsn 300
Db 860 TGTGATACTACGTTCCAGATAAGTTTAAAGGTTTTCATCAACAACAACAAATCGTTGGAAT 919
QY 301 LysSerPheArgGluSerIleLeuSerValLysLysIleMetAsnAsnProPheVal 320
Db 920 AAGTCATTTTATAGGGAGCTATTATCTCTGTTAAGAAGTATTATGCCACCAAGTGT 979
QY 321 AlaLeuTrpThrIleLeuGluValSerMetPheMetMetLeuValTyrSerValValAsp 340
Db 980 GCTGTTTGGACTATACAGAGTTTCCATGTTTCATCATGCTAGTTATTTCATCTTAGC 1039
QY 341 PhePheValGlyAsnValArgGluPheAspTrpLeuArgValLeuAlaPheLeuValIle 360
Db 1040 TTATTTGATGAGAGGCTCAAGAATTTAATCTCATAAACTGGTGTCTTTTATGTTATT 1099
QY 361 IlePheIleValAlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPhe 380
Db 1100 ATTTTCATAGTAGCTCTTGTAGAAATGTTTACATGTTTAAAGCATCCATTTGCTTTT 1159
QY 381 LeuLeuSerProPheTyrGlyValIleHisLeuPheValLeuGlnProLeuLysLeuTyr 400
Db 1160 TTATTTGTCACCGTTTATGGGATGATACATCTATTCGTTTTCACACTCTTAAGATAT 1219
QY 401 SerLeuPheThrIleArgAsnAlaAspTrpGlyThrArgLysLys 415
Db 1220 TCGTTTATTACTATAAGAAATGCTACATGGGGAACCTCGTAAAG 1264

```

RESULT 6
STRHASAO

LOCUS	STRHASO	1350 bp	DNA	linear	BCT 20-APR-1994
DEFINITION	Streptococcus pyogenes WF14 hyaluronate synthase (hasA) gene, complete cds.				
ACCESSION	L21187				
VERSION	L21187.1	GI:469227			
KEYWORDS	hyaluronate synthase.				
SOURCE	Streptococcus pyogenes				
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.				
REFERENCE	1. (bases 1 to 1350)				
AUTHORS	Dougherty, B.A. and van de Rijn, I.				
TITLE	Molecular characterization of hasA from an operon required for hyaluronic acid synthesis in group A streptococci				
JOURNAL	J. Biol. Chem.	269	(1),	169-175	(1994)
MEDLINE	94103204				
PUBMED	8276791				
REFERENCE	2. (bases 1 to 1350)				
AUTHORS	Dougherty, B.A.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-JUL-1993) Brian A. Dougherty, Molecular Biology and Genetics, Johns Hopkins University School of Medicine, 725 N. Wolfe St., Baltimore, MD 21205, USA				
COMMENT	On Apr 8, 1994 this sequence version replaced gi:310625.				
FEATURES	Location/Qualifiers				
source	1..1350				
	/organism="Streptococcus pyogenes"				
	/strain="WF14"				
	/db_xref="taxon:1314"				
-35_signal	2..7				
-10_signal	25..30				
promoter	36				
	/note="transcription start site"				
	/evidence=experimental				
RBS	62..73				
	/note="proposed GTG initiation codon (79..81) downstream from RBS; DNA is complementary to the 3' terminus of streptococcal 16S rRNA"				
gene	79..1338				
	/gene="hasA"				
CDS	79..1338				
	/gene="hasA"				
	/function="required for hyaluronate biosynthesis"				
	/note="GTG start codon"				
	/codon_start=1				
	/transl_except="(pos:79..81,aa:Met)				
	/transl_table=11				
	/product="hyaluronate synthase"				
	/protein_id="AAA17725.1"				
	/db_xref="GI:473747"				
	/translation="MDPKKTLVLSFPLISLILNNVLGTSVGVGYVLLIYLVKIGLSFLYPPFKGPHDYKVAVIETSPNEASLETLKASVAGTTPLETSITVDVGSNTDAQLIEEVKRNDEIVACRVIVHRSVLNKKRHAQAFERSADYFLIVDSITYPNALAEILKSFENDEVTAAAGHIANRQNTLLTLDIRYDNAGFYGRASQSGTGNILVCSGLSIYRREIVTIPMLRYKQTFGLPVSIGDDRLCTNYALDLGRVTQSLTARCDTPFOLSKYLKQQRNKKSEFRESILSKILSNPIVALVTIFEVYVPMMLIYVAGLNLLNQAIQLDLILKLPAPLSITFIIVALCRNHYHMKHPASFLSPLIGILHLFVLPQKLSILCTIKTLTGWETGRKKKTIFF"				
misc_feature	1002				
	/gene="hasA"				
	/note="site of Tn916 insertion in WF62 (hyaluronate synthase mutant)."				
BASE COUNT	427 a	217 c	223 g	483 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	6,85e-133				Length: 1350
Score:	1588.50				Matches: 298
Percent Similarity:	85.44%				Conservative: 54
Best local Similarity:	72.33%				Mismatches: 59
Query Match:	74.06%				Indels: 1
DB:	1				Gaps: 1

```

QY      365 AlaLeuCyAsgAnleHsTyrMetLeuLysHisProLeuSerPheLeuLeuSerPro 384
Db      1168 GCTTATGCTGTAATGTCATATATATGCTCAACATCTGCTAGTGTGTTTTCGTTATCTCT 1227
QY      385 PheTyrGlyValLeuHisLeuPheValLeuClpProLeuLysLeuTyrSerLeuPheThr 404
Db      1228 CTGTATGGAATATTACACTGTGTTCTCTACAGCCCTAAACACTTATTCTTTATGCACC 1287
QY      405 IleArgAsnAlaAspTyrGlyThrArgLysLysLeu 416
Db      1288 ATTAAATACGGAATGGGACACGTAAGAAGTC 1323

RESULT 7
AE010123      10851 bp      DNA      linear      BCT 03-APR-2002
LOCUS      Streptococcus pyogenes strain MGAS8232, section 171 of 173 of the
DEFINITION      complete genome.
ACCESSION      AE010123 AE009949
VERSION      AE010123.1 GI:19749282
KEYWORDS
SOURCE      Streptococcus pyogenes MGAS8232
ORGANISM      Streptococcus pyogenes MGAS8232
Bacteria: Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus
1 (bases 1 to 10851)
Smoot,J.C., Barbano,K.D., Van Gompel,J.J., Smoot,L.M.,
Chaussee,M.S., Sylva,G.L., Sturdevant,D.E., Ricklefs,S.M.,
Porcella,S.F., Parkins,L.D., Beres,S.B., Campbell,D.S., Smith,T.M.,
Zhang,Q., Kapur,V., Daly,J.A., Veasy,L.G. and Musser,J.M.
Genome sequence and comparative microarray analysis of serotype M18
Group A Streptococcus strains associated with acute rheumatic fever
outbreaks
Proc. Natl. Acad. Sci. U.S.A. 99 (7), 4668-4673 (2002)

JOURNAL
MEDLINE      21927593
PUBMED      11917108
REFERENCE
1 (bases 1 to 10851)
Smoot,J.C., Barbano,K.D., Van Gompel,J.J., Smoot,L.M.,
Chaussee,M.S., Sylva,G.L., Sturdevant,D.E., Ricklefs,S.M.,
Porcella,S.F., Parkins,L.D., Beres,S.B., Campbell,D.S., Smith,T.M.,
Zhang,Q., Kapur,V., Daly,J.A., Veasy,L.G. and Musser,J.M.
Direct Submission
Submitted (31-JAN-2002) Laboratory of Human Bacterial
Pathogenesis/Rocky Mountain Laboratories/NIH, 903 S. 4th St.,
Hamilton, MT 59840, USA
FEATURES
source
1. .10851
/organism="Streptococcus pyogenes MGAS8232"
/mol_type="genomic DNA"
/strain="MGAS8232"
/db_xref="taxon:186103"
complement(238..852)
/gene="spym18_2226"
complement(238..852)
/gene="spym18_2226"
/gene="spym18_2226"
/feature="best blast match gb|AAK34820.1| (AE006636)
hypothetical protein [Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAL98659.1"
/db_xref="GI:19749283"
/translation="MFKKENLQRYFNFGVIALTILAIIPAFSSKNADTKSYAKKS
ESKWTIDAKPNHAIYKESKAKSIASPIPTVNSVAPTVTEBPVQVEVQ
TVQVSSVLPNNVNLNGTAGVGSQAQAAATGVPQSTWEHILARESNGPN
AANASGASGLFQTPMGQSTATVEDQVNAALKAYSQAQLSAWGY"
complement(1223..2023)
/gene="spym18_2228"
complement(1223..2023)
/gene="spym18_2228"
/feature="best blast match gb|AAK34821.1| (AE006636)
conserved hypothetical protein [Streptococcus pyogenes M1
GAS]"

```

```

/codon_start=1
/transl_table=1
/product="conserved hypothetical protein"
/protein_id="AAL98660.1"
/db_xref="GI:19749284"
/translation="MKLILGRYIPGDSLHRLDPRSKLLAMIYIYVIFANNVVTN
LMLFTLAVFLSKIKLFFLNGVKPMIGLITLTFOMFSGGKVIYFWMFISIT
DLGLSQALIFMREVLIIFFSTLLATLPLSLDAVSLKPLRKVPKPAHEIGLML
SLSLFVPTLMDYTRIMNAQARGVDFGQNLIOKVKSIIPILIPFASFRKRAL
ATAMEARGYQGGGRKRYQLDWQLKDSLAIGIVSLGLLFFLKTSL"
complement(2016..2858)
/gene="spym18_2229"
complement(2016..2858)
/gene="spym18_2229"
/feature="best blast match gb|AAK34822.1| (AE006636)
putative ABC transporter (ATP-binding protein)
[Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=1
/product="putative ABC transporter (ATP-binding protein)"
/protein_id="AAL98661.1"
/db_xref="GI:19749285"
/translation="MSINLQVSYTYQAGTFPFRGALFNINLIDLDGTYAFIHTGTS
GKSTLMQLLGLHVPTGIVSKODINHSKKEIKSRKHVGLVQFPESQLEET
VLGDAFGQNGVSPERAEALAREKIALVIGISENFERNFELSGGGMVRAIGIL
AMQPKVLIDEPDTAGLDPKRKELMTIPFKLHSGMTIVLTHLMDVANYADFVVL
DKGLTSGSKPKPTIFQOVSLLEBKQLGVKPYTKLAQRLVDRGIPISLPLTLELREY
LRHG"
complement(2834..3724)
/gene="spym18_2230"
complement(2834..3724)
/gene="spym18_2230"
/feature="best blast match gb|AAK61484.1| (AF082738) ABC
transporter (ATP-binding protein) [Streptococcus
pyogenes]"
/codon_start=1
/transl_table=1
/product="ABC transporter (ATP-binding protein)"
/protein_id="AAL98662.1"
/db_xref="GI:19749286"
/translation="MILKEQAFRLRLNMSALIELKKVTFNYHKDQKPTLDGVS
FHYQEWLSIIHGHSKSTIRLIDGLPEPSGIYVDGLTITNWEIHRKIGM
VFQNDPQVATVEDDVAEGLENKGAHEDIKERNVHALELVGMQFKEKEPARLSG
GQKQVATAGVAKMKPRIILDEATSMLDPKRLELEIKIRNRDQQLVIVSITHDL
DEVALSDRVLMKDGQVESTPEQLFARGDELILGLDIPFTTSVVMQLQESGYPID
YGLTEKLENQLQLISKM"
complement(3675..4217)
/gene="spym18_2231"
complement(3675..4217)
/gene="spym18_2231"
/feature="best blast match gb|AAK34824.1| (AE006636)
phosphatidylglycerophosphate synthase [Streptococcus
pyogenes M1 GAS]"
/codon_start=1
/transl_table=1
/product="phosphatidylglycerophosphate synthase"
/protein_id="AAL98663.1"
/db_xref="GI:19749287"
/translation="MIKENIPNLTLVRIAMIPFFLTSSSNKVGWHIFAIVPAI
ASFTDYDLARKVHVSNFGKFDADPLADKMLVMSAFIMLVGLGLVPANVSAVCCR
ELAVTGRLLLVETGGKVLAAAMGLKINTATQMLSIILLCHWIFLGNVLICLALFFT
IVSGDYKPGASFLFKTEK"
complement(4231..5256)
/gene="spym18_2232"
complement(4231..5256)
/gene="spym18_2232"
/feature="best blast match gb|AAK61482.1| (AF082738) unknown
[Streptococcus pyogenes]"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAL98664.1"
/db_xref="GI:19749288"

```



```

/product="hyaluronan synthase"
/protein_id="AAA17982.1"
/db_xref="GI:410046"
/translation="MISILLYNMLFGTSTVGIVGILTYLVIKLGLSLYEPKPGK
NPQDYKVAAPSPVSNEDAESELETLSKVAQTYPYSEIYVDGSSNTDAQLIBERY
NREWDICRNVIHRSVLNVKGKHAQAFAERSADUVFLTVDSDTIYPNALEEELKSF
NDEVTATGHYNNARKNQINLLTRIDIRDNAGFVERAAQSITGNILVCGSPLSIYREV
REVILPNLEGRKNQTFIGLPVSIGDDRLCTNYAIDIGRTVYSTACRCDTLDFPQIAKY
LKQONWNKSFPRESIIYSVKKLISNPYVALWTIEVVWFMMLIVAIGNLFNFQAIQLDIFL
LIKLFALFSIFPIFVALCRNVHYWKHPASFLSPLYLGILHLFVLQPLKLYSLCTIKNT
EMGTFRKKVTIPK"
CDS
178..1380
/gene="hasa"
/notes="possible alternate start codon (TTG) in-frame at
-15; putative"
/codon_start=1
/transl_table=11
/product="hyaluronan synthase"
/protein_id="AAA17983.1"
/db_xref="GI:410045"
/translation="MYILNMYLFGSTVGIVGYVILTYLVIKLGLSFLYEPFKNGHD
YKVAAPISYNDEAESLETLSKVAQTPYSEIYVDGSSNTDAQLIEEYVNRVE
DICRNVIHRSVLNVKGKHAQAFAERSADUVFLTVDSDTIYPNALEEELKSFNRE
VYATGHLNARNQTNLLTRIDIRDNAGFVERAAQSITGNILVCGSPLSIYREVVI
IPNLRYKNOTFLGPLYSIGDDRCCLTNVAIDIGRTVYSTACRCDTVDPFQIAKYLQO
NHNKNSFFRESIIYSVKKLISNPYVALWTIEVVWFMMLIVAIGNLFNFQAIQLDIFL
FAPLSIFIIVALCRNVHYWKHPASFLLSPLYLGILHLFVLQPLKLYSLCTINKTENGT
EKKTVPFK"
CDS
193..1380
/gene="hasa"
/function="biosynthesis of hyaluronan; enzyme"
/notes="putative"
/codon_start=1
/transl_table=11
/product="hyaluronan synthase"
/protein_id="AAA17984.1"
/db_xref="GI:410046"
/translation="MYLFGTSTVGIVGYVILTYLVIKLGLSFLYEPFKGNPHDKYKAA
VIPSURDESLEETLSKVAQTPYSEIYVDGSSNTDAQLIEEYVNRVEDVICRN
VIVHRSVLNVKGKHAQAFAERSADUVFLTVDSDTIYPNALEEELKSFENDETYYAT
GHLNARNQTNLLTRIDIRDNAGFVERAAQSITGNILVCGSPLSIYREVVIIPNL
RYKNOTFLGPLYSIGDDRCCLTNVAIDIGRTVYSTACRCDTVDPFQIAKSYLKQNRNKK
SFFRESIIYSVKKLISNPYVALWTIEVVWFMMLIVAIGNLFNFQAIQLDIFLKFALFS
IIFIVALCRNVHYWKHPASFLLSPLYLGILHLFVLQPLKLYSLCTINKTENGTTRKKYV
IFK"
BASE COUNT      462 a   223 c   236 g   519 t
ORIGIN

Alignment Scores:
Pred. No.:          1,14e-132           Length:          1440
Score:             1586.50              Matches:         298
Percent Similarity: 85.19%               Conservative:    53
Best Local Similarity: 72.33%            Mismatches:     60
Query Match:       73.96%                Indels:         1
DB:                1                    Gaps:          1

US-09-469-200D-2 (1-417) x STRASA (1-1440)

Qy  5 LysAsnLeuUleThrValValAlaIaphSerIlePheTrpValLeuLeuUleThrValAsn  24
||| ::| ||:::||::|||
Db  133 AAAAAAACCTTAATTGTTTTATTCCTTTATTTTTTGATATCATCTGATTTATCTAAT  192
||| ::| ||:::||::|||

Qy  25 ValTyrLeuPheGlyAlaLysGlySerLeuSerIleThrGlyPheLeuLeuUleIleThr  44
::| ::| ||:::||::|||
Db  193 ATGTCATCATTTGGGAACA---TCAACTGTAGGAATTTATGCAGTAAATAATAAACCCT  249
||| ::| ||:::||::|||

Qy  45 LeuLeuValLysMetSerLeuSerPhePheTyrLysProPheLysGlyArgGlaGlyGln  64
||| ::| ||:::||::|||
Db  250 CTAGTTATCAAACTTGATATCTCTTCCTTTATGAGCCATTTAAAGGAATTCACATGAC  309
||| ::| ||:::||::|||

Qy  65 TyrLysValAlaLaIalleIleProSerTyrAsnGluAspAlaGluSerLeuLeuGluThr  84
||| ::| ||:::||::|||
Db  310 TATAAGTTGCTCTGTAATTCCTTTATTAATGAAGATGCCGAGCTATTATTAGAACA  360
||| ::| ||:::||::|||

```


QY 85 LeuLysSerValGlnGlnThrTyrProLeuAlaGluLeuTyrValValAspGly 104
 Db CTTAAAGTGTGTACACAGACCTATCCGTTATCAGAAATTTATATTGTCATGAGG 429
 QY 105 SerAlaAspGluThrGlyLeuArgGluAspTyrValArgAspThrGlyAspLeu 124
 Db AGTTCAACACACATCAATCAATTAATGAGAGTATGTAATAGACAGATGATAT 489
 QY 125 SerSerAsnValIleHisArgSerGluLysAsnGlnGlyLysArgHisAlaGlnAla 144
 Db TGTCCAAACCTTTATCCTCCGTTCCCTTGTCAATAAAGGAAAGCCATCTCAAGCG 549
 QY 145 TrpAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSerAspThrTyrIle 164
 Db TGGGATTTGAAGAATCTGACGCTGAGCTTTTAAACGGTAGACTACATCTATATC 609
 QY 165 TyrProAspAlaLeuGluLeuLysThrPheAsnAspProThrValPheAlaAla 184
 Db TATCCAAATGCCTTAGAGAACTCTTAAAGGCTTCAATGATGACAGCTTTATCTGCA 669
 QY 185 ThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArgLeuThrAspIle 204
 Db ACAGGACATTTGAATCTGCTAGAAACAGACAACTATCTATTAAACGGACTTACAGATATC 729
 QY 205 ArgTyrAspAsnAlaPheGlyValGluArgAlaGlnSerValThrGlyAsnIleLeu 224
 Db CGTTACGATTAATGCCCTTGGGGTGAGCGCTGCTCAATCATTTACAGGTATATTTTA 789
 QY 225 ValCysSerGlyProLeuSerValTyrArgArgGluValValProAsnIleAspArg 244
 Db GTTGTCTCAGGACCATGAGTATTTATCGACGTGAAGTATTTCTTAACCTAGAGCGC 849
 QY 245 TyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAspArgCysLeuThr 264
 Db TATAAATCAACATCTCTAGGTTACCTGTAGCTGGGATGGGATGATCGATGTTTAAACA 909
 QY 265 AsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLysCysIleThrAsp 284
 Db AATTATGCTATTGATTAGGACGACCTGCTACCAATCAACAGCTAGATGCTACTGAT 969
 QY 285 ValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrpAsnLysSerPhePhe 304
 Db GTACCTTCCCAATTAAGATTAATTAAGCAACAAATCGATGSAATAAATCTTTTTT 1029
 QY 305 ArgGluSerIleLeuSerValLysLysIleMetAsnAsnProPheValAlaLeuTrpThr 324
 Db ACAGAACTATATTCTGTATAAATAATCTCTTAATCCATCCATCGTGTGCTATGGACT 1089
 QY 325 IleLeuGluValSerMetPheMetLeuValTyrSerValValAspPheValGly 344
 Db ATTTTCGAAGTCGTTATGTTTATGATGTTGATTCGCAATTTGGGAATCTTTTGTATTA 1149
 QY 345 AsnValArgGluPheAspTrpLeuArgValLeuAlaPheLeuValIleLeuPheVal 364
 Db CAAGCTATCAATAGACCTATTAAACCTTTTTCCTTTTATCCATCATCTTTATCGTT 1209
 QY 365 AlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPheLeuLeuSerPro 384
 Db GCTTTATGCTGTAATGCTTATATGCTCAACATCTGCTAGTATTTTGTATCTCCT 1269
 QY 385 PheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyrSerLeuPheThr 404
 Db CRTGATGAATATTACATGTTGTTGCTTACAGCCCTTAAACCTTTATCTTTATGACACC 1329
 QY 405 IleArgAsnAlaAspTrpGlyThrArgLysLysLeu 416
 Db ATTAATAATAGCAATGGGAACACAGTAAAAAGGTC 1365

RESULT 9
 AR232201
 LOCUS 1512 bp DNA linear PAT 20-DEC-2002
 DEFINITION Sequence 1 from patent US 6455304.

ACCESSION AR232201
 VERSION AR232201.1 GI:27274189
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1512)
 AUTHORS Weigel, P.H., DeAngelis, P.L. and Papaconstantinou, J.
 TITLE Hyaluronate synthase gene and uses thereof
 JOURNAL Patent: US 6455304-A 1 24-SEP-2002;
 FEATURES Location/Qualifiers
 source 1..1512
 BASE COUNT 484 a 227 c 245 g 556 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,22e-132 Length: 1512
 Score: 1586.50 Matches: 298
 Percent Similarity: 85.19% Conservative: 53
 Best Local Similarity: 72.33% Mismatches: 60
 Query Match: 73.96% Indels: 1
 DB: Gaps: 1
 US-09-469-200D-2 (1-417) x AR232201 (1-1512)
 QY 5 LysAsnLeuIleThrValValAlaPheSerIlePheTrpValLeuLeuIleTyrValAsn 24
 Db 205 AAAAAAATCTTAATTTATCTCTTTATTTTGTATATCTATCTGATTTATCTAAAT 264
 QY 25 ValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeuLeuIleAlaTyr 44
 Db 265 ATGTATCTATTGGACA---TCACCTGTAGGAATTTATGGAGTAATATTAATACCAT 321
 QY 45 LeuLeuValLysMetSerLeuSerPhePheTyrLysProPheLysGlyArgAlaGlyGln 64
 Db 322 CTAGTTTCAAACTTGGATATCTTTCTCTTATGAGCCATTTAAAGGAATCCACATGAC 381
 QY 65 TyrLysValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSerLeuLeuGluThr 84
 Db 382 TATAAAGTTGCTGTGTAATTTCTTATTAATGAAGATGCGGAGTCATTTATAGAAACA 441
 QY 85 LeuLysSerValGlnGlnThrTyrProLeuAlaGluIleTyrValValAspAspGly 104
 Db 442 CTTAAAGTGTGTACACAGACCTATCCGTTATCAGAAATTTATTTGTTGATGAGGG 501
 QY 105 SerAlaAspGluThrGlyLysArgIleGluAspTyrValArgAspThrGlyAspLeu 124
 Db 502 AGTTCAACACACAGATGCAATACAAATTAATTCAGAGATGTAATAAGAGTGGATAT 561
 QY 125 SerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArgHisAlaGlnAla 144
 Db 562 TGTGAAACGTTATCGTTTACCGCTTCCCTTGTCAATAAAGGAAACGCCATGCTCAAGCG 621
 QY 145 TrpAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSerAspThrTyrIle 164
 Db 622 TGGGATTTGAAGAATCTGACGCTGACGCTTTTAAACGGTAGACTACATCTATATC 681
 QY 165 TyrProAspAlaLeuGluGluLeuLysThrPheAsnAspProThrValPheAlaAla 184
 Db 682 TATCCAAATGCCCTTAGAAGAACTCTCTAAAGGCTTCAATGATGACAGCTTTATGCTGCA 741
 QY 185 ThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArgLeuThrAspIle 204
 Db 742 ACAGGACATTTGAATGCTAGAAACAGACAACTATCTATTAAACGGACTTACAGATATC 801
 QY 205 ArgTyrAspAsnAlaPheGlyValGluArgAlaGlnSerValThrGlyAsnIleLeu 224
 Db 802 CGTTACGATTAATGCCCTTGGGGTGAGCGCTGCTCAATCATTTACAGGTATATTTTA 861
 QY 225 ValCysSerGlyProLeuSerValTyrArgArgGluValValProAsnIleAspArg 244
 Db 862 GTTTGCTCAGGACCATGAGTATTTATCGAGCTGAAGTATTTCTTAACCTAGAGCGC 921

QY	245	TyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspArgCysLeuThr	264	gene	/db_xref="taxon:160490"
Db	922	TATAAAATCAAACTTCAGTTACCTAGTTAGCTAGGATCGATGTTTAAACA	981	CDS	246..1505 /gene="haha" /note="synonym: SPY2200"
QY	265	AsnTyAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLysCysIleThrAsp	284		246..1505 /gene="haha"
Db	982	AATYATGCTATGATTAGGACGCATGCTACCAATCACAGCTAGATGTGATCAGAT	1041		/note="Best Blast hit = pir A53100 hyaluronate synthase A (Hsa) - Streptococcus sp. (group A) >gi1473747 gb AAAI7725.1 (L21187) hyaluronate synthase [Streptococcus pyogenes]"
QY	285	ValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTirpAsnLysSerPhePhe	304		/codon_start=1
Db	1042	GTACCTTCCCAATTAATAAGTATTATTAAGCAACAATCGATGGAATAATCTTTTTT	1101		/transl_table=11
QY	305	ArgGluSerIleLeuSerValLysIleMetAsnAsnProPheValAlaLeuTrpThr	324		/product="hyaluronate synthase"
Db	1102	AGAAATCTATTAATTCGTGTAATAAATAATCTTTTAAATCCATCGCTTGCTATGACAT	1161		/protein_id="AAK34828.1"
QY	325	IleLeuGluValSerMetPheMetMetLeuValTyrSerValValAspPhePheValGly	344		/db_xref="GI:13623173"
Db	1162	ATTTCGAAGTCGTATGTTATGATGCTGATGTCGCAATCGGAATCTTTGTTTAAT	1221		/translation="MPFKKTLVLISFLISLILYLMVLTGTVGVIGVILITVLKGLSPLEPKPGPHDYKVAVIPSYNEDEASLETKLSVLAQTYPLSEIYVDDGSNDLAIQLEEVYVNRVDICRNVIYHRSLYNKRKAQWAPERSADVFLIVDDTYIYNALBELLSFNDEFVTAATGHLNARNKQTNLLTRLDIRYDNAPGVFAAOSLTGNLISGGLISYIRREVILINLRYKQTFGLSPVSGDRCITNTYALIDGRTVOSCARCTDVPGLSKYLKQONRNKSFKEKLSISVKKLLSNPVALMTIFEVYPMMLIYVAGNLLENQAIQDLIKLFLPSLIIIFIVALCRNVHYMKHPASFLSPGLYGLHLFVLOPLKLYSLICFTIKNTWGTTRKKVIETK"
QY	345	AsnValArgGluPheAspTrpLeuArgValLeuAlaPheLeuValIlePheIleVal	364	gene	1541..2749 /gene="hasb"
Db	1222	CAAGCTATTCAAATAGACCTTATTAACACTTTTTCCTTTATCCATCATCTTATCGTT	1281		/note="hasb"
QY	365	AlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPheLeuLeuSerPro	384	CDS	/note="synonym: SPY2201"
Db	1282	GCTTATGCTGATGTTATGATGCTATGATGCTGATGCTGATGCTGATGCTGATGCT	1341		1541..2749 /gene="hasb"
QY	385	PheTyGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyrSerLeuPheThr	404		/EC_number="1.1.1.22"
Db	1342	CTGTATGGAATATTACACTTGTGCTTACAGCCCTTAAACTTATTCCTTATGACCC	1401		/note="Best Blast hit = sp O07172 UDG_STRPY UDP-GLUCOSE 6-DEHYDROGENASE (UDP-GLC DEHYDROGENASE) (UDP-GLCDH) (UDPGDH) >gi11075776 pir A46089 UDPglucose 6-dehydrogenase (EC 1.1.1.22) - Streptococcus pyogenes >gi18569328 pdb 1DLIIA Chain A, The First Structure Of Udp-Glucose Dehydrogenase (Udpqdh) Reveals The Catalytic Residues Necessary For the Two-Told Oxidation >gi12951891 gb AAA26899.1 (L08444) UDP-glucose hydrogensease [Streptococcus pyogenes]"
QY	405	IleArgAsnAlaAspTrpGlyThrArgGlyLysLeu	416		/codon_start=1
Db	1402	ATTAAATAATACGAATCGGAACACGATAAAGGTC	1437		/product="UDP-glucose 6-dehydrogenase"
LOCUS	AE006637	9844 bp	DNA	linear	/protein_id="AAK34829.1"
DEFINITION	Streptococcus pyogenes M1 GAS strain SF70, section 166 of 167 of the complete genome.				/db_xref="GI:13623174"
VERSION	AE006637 AE004092				/translation="MKIAYAGSGVGLSLQNEVTIVDILPSKVDKINGLSPIQDEYIYLKSLQSLKATLDSKAKEALVIATITNTNSRINFTYDTHQVETVIEVLSNSHATLIITKSTPIGTFITETMRQFTDRIIFSEPIRESKALDNIYLPRIIVSCENDSPKADAEKFAALLKSAKNPNVPLINGSEAEAVKLFANTYALALPRAVYNELDYIAERKLSNHHMIQIGSYDORIGMHNPSFGYGGVCLPKDTKOLLANTNNIPOTLIEAIVSSNNVRKSYIAKQIINVLEERSPVGVYRLIMKSNDSNFESALKDVIDLKSKDKITILIVEMLNKLESDQSVLVNDLENFKQANILVTRNYDNLQDVKNKYVSRDIFNRD"
KEYWORDS	Streptococcus pyogenes M1 GAS				2931..3845 /gene="hasc"
SOURCE	Streptococcus pyogenes M1 GAS				/note="synonym: SPY2202"
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.				2931..3845 /gene="hasc"
REFERENCE	1 (bases 1 to 9844)				/EC_number="2.7.7.9"
AUTHORS	Ferretti,J.J., McShan,W.M., Adjic,D., Savic,D., Savic,G., Lyon,K., Primeaux,C., Sezate,S.S., Surorov,A.N., Kenton,S., Lai,H., Lin,S., Qian,Y., Jia,H.G., Najjar,F.Z., Ren,Q., Zhu,H., Song,L., White,J., Yuan,X., Clifton,S.W., Roe,B.A. and McLaughlin,R.E.				/note="Best Blast hit = sp Q54713 HASC_STRPY UDP-GLUCOSE-1-PHOSPHATE URIDYLTRANSFERASE (UDP-GLUCOSE PHOSPHORYLASE) (UDPGP) (ALPHA-D-GLUCOSYL-1-PHOSPHATE URIDYLTRANSFERASE) (URIDINE DIPHOSPHOGLUCOSE PHOSPHORYLASE) >gi1995767 gb AAA91810.1 (U33452) UDP-glucose pyrophosphorylase [Streptococcus pyogenes]"
TITLE	Complete genome sequence of an M1 strain of Streptococcus pyogenes				/codon_start=1
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 98 (8), 4658-4663 (2001)				/product="UDP-glucose pyrophosphorylase"
MEDLINE	21192684				/protein_id="AAK34830.1"
PUBMED	11296296				/db_xref="GI:13623175"
REFERENCE	2 (bases 1 to 9844)				/translation="MTKVRKALIPAGLGTRFLPATYALAKEMLPVDPKPTQIETVEALKSIEEILVVTGKAKSIEDHDSNFEYLNQAKGNELLKLVDETAINHTIRI QSHPRGLDVAQKAFVGNPEFFVNLGDMDLITNASAKPLTKGLMEDYDTHASTIAYMKVPYHEDYSYGVIAQPKAVGLYSVDTEVEKPOPEADPSDLAIGRYLTPRIIFGILEQTPGAGNEVOLTAIDTIAKTORVFAEEKNGRYDVGDKFGFKTSTIDYALFIV
AUTHORS	Ferretti,J.J., McShan,W.M., Adjic,D., Savic,D., Savic,G., Lyon,K., Primeaux,C., Sezate,S.S., Surorov,A.N., Kenton,S., Lai,H., Lin,S., Qian,Y., Jia,H.G., Najjar,F.Z., Ren,Q., Zhu,H., Song,L., White,J., Yuan,X., Clifton,S.W., Roe,B.A. and McLaughlin,R.E.				
TITLE	Direct Submission				
JOURNAL	Submitted (10-APR-2001) Department of Microbiology and Immunology, University of Oklahoma Health Sciences Center, 940 SL Young Blvd, Oklahoma City, OK 73104, USA				
FEATURES	Location/Qualifiers				
source	1..9844				
	/organism="Streptococcus pyogenes M1 GAS"				
	/mol_type="genomic DNA"				
	/strain="SF370"				
	/serotype="M1"				

Db	855	CGTTACGATAATGCTTTGGGTGGAGCGTCTGCTCAATCAATTAACAGGTAATATTTTA	914	1.	53896	/organism="Streptococcus pyogenes MGAS315"
QY	225	ValCysSerGlyProLeuSerValTyrrArgGluValValProAsnIleAspArg	244			/mol_type="genomic DNA"
Db	915	GIITGTCAGGACCATTTAGTATTATTCAGCGTGAAGTATTATTCCTAAGTACGAGCG	974			/strain="MGAS315"
QY	245	TVrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspArgCysLeuThr	264			/serotype="M3"
Db	975	TATAAATCAAAATTCCTAGTTCCTAGCTTAGCATTTGGGATGATCGATTTTAA	1034			/db_xref="taxon:198466"
QY	265	AsnTyraIleAspLeuGlyLysThrValTyrrGlnSerThrAlaLysCysIleThrAsp	284			/note="group: A"
Db	1035	AATTATGCTATTGATTATAGGACGCACTGCTACCAATCAACAGCTAGATGTGATCAT	1094			/complement(200..1480)
QY	285	ValProAspLysMetSerThrTyrrLeuLysGlnGlnAsnArgTrpAsnLysSerPhePhe	304			/gene="hlss"
Db	1095	GTACCTTCCCAATTAAGATTTATTAAGCAACAATCGATGAATTAATCTTTT	1154			/note="synonym: SpyM3_1815"
QY	305	ArgGluSerIleIleSerValLysLysIleMetAsnAsnPropheValaIleLeuTrpThr	324			/complement(200..1480)
Db	1155	AAAGATCTATTATTTCTGTTAAATAATCTTTCTAATCCCATCGTTCCTTATGGACT	1214			/gene="hlss"
QY	325	IleLeuGluValSerMetPheMetMetLeuValTyrrSerValValAspPheValGly	344			/note="synonym: SpyM3_1816"
Db	1215	ATTTCGAAGTCGTTATGTTTATGATGTTGATTCGCAATTCGGAATCTTTTCTTAAT	1274			/protein_id="AAM80422.1"
QY	345	AsnValArgGluPheAspTrpLeuArgValLeuAlaPheLeuValIleIlePheIleVal	364			/db_xref="GI:21905567"
Db	1275	CAAGCTATTCAATAGACCTTATTAAGTCTTTTTCCTTTTATCCATCACTTATCGTT	1334			/translation="MKLOPKGTDLLPQDAKQWQVESVARDTFSOVNYGEIRTPMF
QY	365	AlaLeuCysArgAsnIleHisTyrrMetLeuLysHisProLeuSerPheLeuLeuSerPro	384			EHEVISRSVGDITDIYTKEMDFDKGRHILIRPEGAPVRSYVENKLFAPVQK
Db	1335	GCATTATGTCGTAATGTTTATATATGATCAACATCCTGCTAGTTTTTTTATCTCCT	1394			PVLYLTIGSMERYERPOAGRLREFHOIGVECFGAANPAIDVETIAMAYHFEKLGKO
QY	385	PheTyrrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyrrSerLeuPheThr	404			VTUHLNSPESRISAYRQALIDYITPMRDQSKDSQRDLNPLVLDSEKEDKLA
Db	1395	CTATGGAATGATACACTGTTGTGTCACGCCCTAAACCTTATCTTTATGCACCC	1454			VERAPSTILDYDESSQAFVAKMLEALDIPYVIDTNMVRGLDYNHITFEITISVE
QY	405	IleArgAsnAlaAspTrpGlyThrArgLysLysLeu	416			GSDLTTCAGRDYDSLVGYFGGPTPGFGLGERILMWTEKOGITLPIETEMDVLA
Db	1455	ATTAAATATCGGAATGGGACACGCTAAAGGTC	1490			VLGDGANSKALELVAIRROGTTAERDYLGRKIKAKPKSADTFKAKLVMLGSEVEA
RESULT 11						GKAVIKNESRQEVESFEDMTNFANISEQLLS"
AE014172						1700..1882
LOCUS						/gene="rpmf"
DEFINITION						/note="synonym: SpyM3_1816"
ACCESSION						1700..1882
VERSION						/gene="rpmf"
KEYWORDS						/note="best blastp hit: gb AAL98635.1 (AE010120) 50S
SOURCE						ribosomal protein L32 [Streptococcus pyogenes MGAS8232]"
ORGANISM						ribosomal protein L32
REFERENCE						/codon_start=1
AUTHORS						/transl_table=11
						/product="50S ribosomal protein L32"
						/protein_id="AAM80423.1"
						/db_xref="GI:21905568"
						/translation="MAVPAHTSKAKNKRTHYKLTAPSVQFDETTGDSRSRHSYSL
						KGYKGRKAKANEAK"
						1898..2047
						/gene="rpmg"
						/note="synonym: SpyM3_1817"
						1898..2047
						/gene="rpmg"
						/note="best blastp hit: gb BAK34796.1 (AE006634) 50S
						ribosomal protein L33 [Streptococcus pyogenes M1 GAS], and
						gb AAL98636.1 (AE010120) 50S ribosomal protein L33
						[Streptococcus pyogenes MGAS8232]"
						/codon_start=1
						/transl_table=11
						/product="50S ribosomal protein L33"
						/protein_id="AAM80424.1"
						/db_xref="GI:21905569"
						/translation="MRVNITLKHESGERLYLTSKNKNRPDRILQKKYSPKLRKHYT
						FTEVK"
						2328..2942
						/gene="cadD"
						/note="synonym: SpyM3_1818"
						2328..2942
						/gene="cadD"
						/note="best blastp hit: gb AAK34797.1 (AE006634) putative
						cadmium resistance protein [Streptococcus pyogenes M1
						GAS]"
						/codon_start=1
						/transl_table=11
						/product="putative cadmium resistance protein"
						/protein_id="AAM80425.1"
						/db_xref="GI:21905570"
						/translation="MIONVYTSILYSGTAVDLLIILMLFFAKRSKREDIINLYGOF
						LGWSLILSLIFAFVDYIPSKELIGLLGLPIFGLKVLGDSGDEAIKAEKLSK
						DNKNLIFLIVAMITFASCGADNIGVFVPFTTILNLIANLIVALLTFLVMIYLLVLSAQL

```

gene
AQVPSVGETLEKYSRWFIANVYVLGLGMYILLIENNSFDMLWAVLG"
2954. .3292
/ gene="cadx"
/ note="synonym: SpyM3_1819"
CDS
2954. .3292
/ gene="cadx"
/ note="best blastp hit: gb|AAK34798.1| (AE006634) putative
cadmium efflux system accessory protein [Streptococcus
pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="putative cadmium efflux system accessory
protein"
/ protein_id="AAM80426.1"
/ db_xref="GI:21905571"
/ translation="MKDSCQGVINQNVYATNATNLEKEKVKSLKILSKFTDNKQ
INIFILVAEELCVCDIACLLNSMASHLRLKLANILDTREGKIIYFIKDE
EIRDFNFOLG"
3343. .4266
/ gene="SpyM3_1820"
3343. .4266
/ gene="SpyM3_1820"
/ note="best blastp hit: gb|AAK34799.1| (AE006634)
hypothetical protein [Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/ protein_id="AAM80427.1"
/ db_xref="GI:21905572"
/ translation="MKFEENYSQEIPTRIKRLKKNITQSELNAGQVQVESGRK
PITSVLLNALTAASYTYVFEGLDEFIENLFYFFSILYRLEAVDENLYSFM
SDIISQSLCLIAKTFANFIQKRFMISTETEMDFHKDDIDVWGGSYNPARS
FRNPINELTVIDREMAIDILLMLRNLIRSFNVCNLTFLDKNGEPTFLNDKI
DSINIKWSENSTVETPLNIKILRENPLNPGWVNDILERMKYEDIPKSLYLSVPL
VISKARSILSLKNSGSKDKSKFAQLSL"
4332. .5084
/ gene="SpyM3_1821"
4332. .5084
/ gene="SpyM3_1821"
/ note="best non-GAS blastp hit: gb|AAK79835.1|AE007695_8
(AE007695) FlsK-like DNA segregation ATPase, YDCQ
B. subtilis ortholog [Clostridium acetobutylicum]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/ protein_id="AAM80428.1"
/ db_xref="GI:21905573"
/ translation="MAIPDLDFESSNLOQGLPIVKHFLNLDKSHWAIAGNSGSK
PYALTYFLSVLPMSEGLIIDPKFPTPSQWAKENKIAVIHPVENHKSDFYSQVNEQL
NOCATLQKRAILYDNPHTFLTVIDEVLAISEGVNKNIKKAFFSLISLALIG
HATKILFLGSGQRFQDHTIPISVREQLNVLQIGNINOKTQPLFPDLDPBGIVPTG
HGTGILQVVDNEHSYQVLLPLLCPTYYTKRGIL"
5081. .5683
/ gene="SpyM3_1822"
5081. .5683
/ gene="SpyM3_1822"
/ note="best blastp hit: gb|AAK34801.1| (AE006634)
hypothetical protein [Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/ protein_id="AAM80429.1"
/ db_xref="GI:21905574"
/ translation="MTTKTYFMRSFNLILYLLLIAPICYVFGMLFNLYSVOLSTVTFV
NLDTVKNWDTLSSVCYGFALISLILALISGLMKIQLRNTDLSGNWCOSIYRTFLER
FLFQREKQVSNFQRVTTINPIINGFNRAVRKCIIVIRKDALITVELKVPQQGQ
KILKDTGVLKEEVIQHTDIFFSSPVRVRLNOLMFIGQR"
complement(6159. .6323)
/ gene="SpyM3_1823"
complement(6159. .6323)
/ gene="SpyM3_1823"
/ note="best blastp hit: gb|AAK34556.1| (AE006610)
hypothetical protein [Streptococcus pyogenes M1 GAS]"

```

```

/codon_start=1
/transl_table=11
/product="hypothetical protein"
/ protein_id="AAM80430.1"
/ db_xref="GI:21905575"
/ translation="MLFGTLLIAISYIPMSLWGMWGPVAILWLIIMRSALRIERF
KKTVDVQTY"
complement(6651. .6809)
/ gene="SpyM3_1824"
complement(6651. .6809)
/ gene="SpyM3_1824"
/ note="best blastp hit: emb|CAB65345.1| (Y18538) PR7
protein [Plasmodium falciparum]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/ protein_id="AAM80431.1"
/ db_xref="GI:21905576"
/ translation="MFLINNTFNFEDIFLKHVFVKSSLIIVKVNKIYCKCKPYDNY
QKDKRKA"
6894. .7736
/ gene="SpyM3_1825"
6894. .7736
/ gene="SpyM3_1825"
/ note="best non-GAS blastp hit: gb|AAK99693.1| (AE008462)

Alignment Scores:
Pred. No.: 2,59e-130 Length: 53896
Score: 1585.50 Matches: 298
Percent Similarity: 85.19% Conservative: 53
Best Local Similarity: 72.33% Mismatches: 60
Query Match: 73.92% Indels: 1
DB: 1 Gaps: 1

US-09-469-200D-2 (1-417) x AE014172 (1-53896)

Qy 5 LysAsnLeuIleThrValValAlaPheSerIlePheTrpValLeuLeuIleTyrValAsn 24
Db 33885 AAAAACAATTAAATGGTTTATTCCTTTATGTTGATATCTATCTGATTATCAAT 33944
Qy 25 ValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeuLeuIleAlaTyr 44
Db 33945 ATGTATCTATTGGGAACA---TCAACTGTAGGAATTTATGGAGTAAATATAAACCCTAT 34001
Qy 45 LeuLeuValLysMetSerLeuSerPhePheTyrLysProPheLysGlyValGlaGlyGln 64
Db 34002 CTAGTTATTAACCTGGATATCTCTTCCTTATGAGCCATTAAAGAAATCCACATGAC 34061
Qy 65 TyrLysValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSerLeuLeuIleThr 84
Db 34062 TATAAAGTTGCTGCTGTAATTCCTTCTTATAATGAAGATCCCGAGTCATATTAGAAACA 34121
Qy 85 LeuLysSerValGlnGlnGlnThrTyrProLeuAlaGluIleTyrValValAspGly 104
Db 34122 CTTAAAAGTGTGTTAGCACAGACCTATCCGTTATCAGAAATTTATATTGTTGATGATGG 34181
Qy 105 SerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgAspThrGlyAspLeu 124
Db 34182 AGTTCACACACAGATGCAATACAAATTAATGAGATGTGTAATAGAGATGGGATATT 34241
Qy 125 SerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArgHisAlaGlnAla 144
Db 34242 TGTGAAAGCTTATGTTCCCTTCCCTTGTCAATAAAGAAACGCCATGCTCAAGG 34301
Qy 145 TrpAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSerAspThrTyrIle 164
Db 34302 TGGGCATTTGAAAGATCTGACGCTGACGCTTTTAAACCGTAGATTACAGATTATATC 34361
Qy 165 TyrProAspAlaLeuGluLeuLeuLysThrPheAsnAspProThrValPheAlaAla 184
Db 34362 TATCCAAATGCCCTTAGAAGAACTCTCCAAAAGCTTCAATGATGACAGACGTTATGCTGA 34421
Qy 185 ThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArgLeuThrAspIle 204

```

```

|||||
Db 34422 ACAGGACATTTGATGCTAGAAACAGACAACTAACTATTATPACGGCAGCTTACAGATATC 34481
QY 205 ArgTyrAspAsnAlaPheGlyValGluArgAlaGlnSerValThrGlyAsnIleLeu 224
Db 34482 CGTTAGCATATGCTTTGGGGTGGAGCGTCTGCTCAATCATTAACAGGTAATATTTA 34541
QY 225 ValCysSerGlyProLeuSerValTyrArgArgGluValValProAsnIleAspArg 244
Db 34542 GTTTGCTCAGGACCATTTGATATTATCGAGTGAAGTGAATTATTCCTTACTTAGAGCGC 34601
QY 245 TyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspArgCysLeuThr 264
Db 34602 TATAAAATCAACATTCCTAGTTTACCTTTAGCATTTGGGATGATCGATGTTTAAACA 34661
QY 265 AsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLysCysIleThrAsp 284
Db 34662 AATATGCTATTATTTAGGAGCGACTGCTACCAATCAACAGCTAGATGTGATACATG 34721
QY 285 ValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTyrAsnLysSerPhePhe 304
Db 34722 GTACCTTTCCAAATTAAGATTAATTAAGCAACAATAATCGATGAATTAATCTTTT 34781
QY 305 ArgGluSerIleLeuSerValLysLysIleMetAsnAsnProPheValAlaLeuThrThr 324
Db 34782 AGAGAATCTATTATTTCTGCAAAAAAATCTTTCTAATCCCATCGTTGCTTATGGACT 34841
QY 325 IleGluGluValSerMetPheMetMetLeuValTyrSerValValAspPheValGly 344
Db 34842 ATTTTCGAAGTCGTTATGTTATGATGTTGATGTCGAATTCGGAATCTTTTGTAAAT 34901
QY 345 AsnValArgGluPheAspTyrLeuArgValLeuAlaPheLeuValIlePheIleVal 364
Db 34902 CAAGCTATTCAATAGACCTTATAAAGCTTTTTCCTTTTATCCATCATCTTATCGTT 34961
QY 365 AlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPheLeuLeuSerPro 384
Db 34962 GCTTTATGTCGTAATGTTTATGATGTTTATGTCGAACATCTGCTAGTTTTCCTATCCT 35021
QY 385 PheTyrGlyValIleHisLeuPheValLeuGlnProLeuLysLeuTyrSerLeuPheThr 404
Db 35022 CTGTATGGAATTAACATCTGTTGTCCTCAGCCCTCAAACTTATCTTATGACCC 35081
QY 405 IleArgAsnAlaAspTyrGlyThrArgLysLysLeu 416
Db 35082 ATTAAAAACGAATGGGACACGCTAAAGGCTC 35117

RESULT 12
AP005146 323825 bp DNA linear BCT 13-JUN-2003
LOCUS Streptococcus pyogenes SSI-1 DNA, complete genome, section 6/6.
ACCESSION AP005146 BA000034
VERSION AP005146.1 GI:28811737
KEYWORDS Streptococcus pyogenes SSI-1
SOURCE Streptococcus pyogenes SSI-1
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.

REFERENCE
1
AUTHORS Nakagawa, I., Kurokawa, K., Yamashita, A., Nakata, M., Tomiyasu, Y.,
Okahashi, N., Kawabata, S., Yamazaki, K., Shiba, T., Yasunaga, T.,
Hayashi, H., Hattori, M. and Hamada, S.
TITLE Genome Sequence of an M3 Strain of Streptococcus pyogenes Reveals a
Large-Scale Genomic Rearrangement in Invasive Strains and New
Insights into Phage Evolution
JOURNAL Genome Res. 13 (6), 1042-1055 (2003)
MEDLINE 22683278
REFERENCE 2 (bases 1 to 323825)
AUTHORS Yamashita, A., Nakagawa, I., Kurokawa, K., Nakata, M., Tomiyasu, Y.,
Yamazaki, K., Okahashi, N., Kawabata, S., Yasunaga, T., Hattori, M.,
Hayashi, H. and Hamada, S.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2002) Ken Kurokawa, Osaka University, Genome

```

```

Information Research Center: 3-1, Yamadaoka, Suita, Osaka 565-0871,
Japan (E-mail:ken@gen-info.osaka-u.ac.jp; URL:http://www.gen-info.osaka-u.ac.jp/, Tel:81-6-6879-8365,
Fax:81-6-6879-2047)
genome project
This clone was isolated from a patient presenting with toxic shock
like syndrome.

COMMENT
FEATURES
source
1. 323825
Location/Qualifiers
/organism="Streptococcus pyogenes SSI-1"
/mol_type="genomic DNA"
/strain="SSI-1"
/db_xref="taxon:193567"
complement(422..1012)
/gene="SPS1572"
complement(422..1012)
/gene="SPS1572"
/notes="similar to GB:AAK33432.1 (AE006502) percent
identity 100 in 196 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative ATP-dependent protease proteolytic
subunit"
/protein_id="BAC64667.1"
/db_xref="GI:28811738"
/translation="MIPVIEQTSRGERSYDIYSRLKKDRILMLTGPVEONMANSVIA
QLFLDAQNTKDIYLVNTPGSGVSAGLAIVDMNFIRADYQTIYMGMAASGTVIA
SSGTKGRMLPNAYMIHQPMTGGGTQOTDMAAEEHLKTHRLEKILQAONAGK
TIQIHKAERDYWMSAEETLAYGFIDSEIMENNELK"
complement(1235..1864)
/gene="SPS1573"
complement(1235..1864)
/gene="SPS1573"
/notes="similar to GB:AAK33430.1 (AE006502) percent
identity 100 in 209 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative uracil phosphoribosyltransferase"
/protein_id="BAC64668.1"
/db_xref="GI:28811739"
/translation="MGKQVISHPLIQHKLILRRQTTSTKDFRELVNEIAMLMGYEV
SRDLPLEVDIQTIVSKTVOKLAGKLAIVPILRAGIGMVDGLLSIVPAKVGHTGM
YRNEETLEPVEVLKLPEDINORQIFLVDPLMATGGSATLAVDSLKKRGAANKFVCL
VAAPGVKKLQEAHPDIDIFTAALDHLNHEGYIVPGLGDAGDRLEFGTK"
complement(2032..3666)
/gene="SPS1574"
complement(2032..3666)
/gene="SPS1574"
/notes="similar to GB:AAK33429.1 (AE006501) percent
identity 99 in 544 aa"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC64669.1"
/db_xref="GI:28811740"
/translation="MSTKKQLTQELVQGAHWSTAGNISRLGLGVLIIPWYIMWG
QYALQANFLNNGNYVAYFELLSTTGLNVAIAQVAKYNSMGQTEHSYQLIRSLTKL
MLGLIFSAIMYGLSPFLASGSDDTLVPMHSLSLAVFFPVNSYIRGIFQGHNN
IKPYAISOAEOLIRVIMLLTFLIMKLGSDYASAVTQSTFAAFIGWASMGVLGY
YLMKQGLAAIFSRPMDTVIDIKGLLLETLSKISIPFIVTSAIQAFOLIDQWTFYNT
MTLFTDYSSQLLVIFGYFNANPAKITWVLIANAASIGGVGIALITENYVKDKMAA
RLINNIEMLYNFPALTGAIIILAPRPLYSVFYGCASDRALHFLVAVLFTQLMALYT
LFSPMLQALFENRRAYIFAYGILKVLQIPLIYLHAYGLPALTIALIVVPVILMY
RLYQVTHENRKLKQLLETLLTLMGLVFFVWMLLYAFKPTGSLTSLTLLII
GGGMVYVYATLTLLTHQLDKLIGSKASRLQKLGWH"
3754..5199
/gene="SPS1575"
3754..5199
/gene="SPS1575"
/notes="similar to GB:AAK33428.1 (AE006501) percent

```

```

identity 99 in 481 aa"
/codon_start=1
/evidence-not_experimental
/transl_table=11
/product="putative UDP-N-acetyluramoylalanyl-D-glutamyl-2
, 6-diaminopimelate ligase"
/protein_id="BAC64670.1"
/db_xref="GI:28811741"
/translation="MTTEQLDILKKDHNFREVLDADEVHVGQSFPERLSVDSQC
VDGKLEFAKGAFTKADYIKKATNGSLQIYISEVDYELGIPVVLVIDDKKAMSLIAMA
FYGPQKLLAFTAGTKITTAIFAYHMLKSEKPFAMESTNITLDGKTFKFSQIT
TPESLDLFAMAECTNGMTHLINEVSSQAYLVDRVYGLTFDFVGNLSPDHGPIE
HPTTFDYHKRLMNSRAVVISNSVDHFSFLADQVADQEHFYGLPSDQNIITSOA
FSFAKQLAGHYDIQLIGHNQENAMAGLACLRLGLASLADIKGIAKTRVPRGVH
LTMNHAKVVDYAHNGDSLEKLSVVEEHQTKLMLILGAPGNKGSRRADFGRVH
QHPNLVTLDADPNFDPEDISKETASHIAPPELVISDREQALQKMSLQCGAKDAV
IIACKGADAYQIVKQVAYAGDLATAKHVL"
5445..6227
/gene="Sps1576"
5445..6227
/gene="Sps1576"
/notes="similar to GB:AAK33427.1 (AE006501) percent
identity 99 in 260 aa"
/codon_start=1
/evidence-not_experimental
/transl_table=11
/product="putative ferrichrome ABC transporter
(ATP-binding protein)"
/protein_id="BAC64671.1"
/db_xref="GI:28811742"
/translation="MTTISAEEDTIAYEQRTIIDKLSFYIPEKTIITIGANGCKGS
LLKALLRLPPKQGVYVNGQNTATLEKVAKILALIPQVQATNGITVYELVSYGR
PFHGYENLSPADKKAJHMAQATNVMAYADQPDVALSGQQRVYWLAMALAGOTDT
IFLDPTTYLDLNLHOLEILELVKSLNKDAGKTIIMVLDLNLRSARYSDHLTAMKHGI
HYTGTIADYMTSPIQDIFQIKPVLVDDPIHNCPIVLYTQLQ"
6253..7185
/gene="Sps1577"
6253..7185
/gene="Sps1577"
/notes="similar to GB:AAL97176.1 (AB009984) percent
identity 99 in 310 aa"
/codon_start=1
/evidence-not_experimental
/transl_table=11
/product="putative ferrichrome ABC transporter
(ATP-binding protein)"
/protein_id="BAC64672.1"
/db_xref="GI:28811743"
/translation="MKKLTLLTLLTTLTIACGNQATNHSNTASKSLSPMPQIAGV
TYIGDPKQKRVVSLASTYTGKLLDMNLVGYTSYDKRNPIILAKTVKAKQVAATD
LEATITLPLVIGSTEENIKQAEAPVVISIEYKRDYLVLSDFGRFNKRGAK
KWLKDKTKTAAEYKVAETGDKATFTIMGLYKDYLFEGKDWGRGGEIHQAFHYD
APERKVEYFKQGYLSLSQEVLPDYIGDYVVAEDDGTGSALYESKLWQSIAPVKKH
HVIKVNANVFYFDPLSLEYQLETLEAILLSEN"
7196..8227
/gene="Sps1578"
7196..8227
/gene="Sps1578"
/notes="similar to GB:AAL97175.1 (AB009984) percent
identity 99 in 343 aa"
/codon_start=1
/evidence-not_experimental
/transl_table=11
/product="putative ferrichrome ABC transporter (permease)"
/protein_id="BAC64673.1"
/db_xref="GI:28811744"
/translation="MCLERFCLKKTQIITKSHFWLVFVLISFLCVAYICHLRFQAV
ALSHQDLNSILFKQNGKHANVILLALPLRFLGATLGTCSALAVSGTIMGATNPPIAE
PELLGNAGALVALVAFPHLHYSLTLLSLIGSLAATLVFLGSLYQSGKGYHOL
RLVLGAWGILLSALGQGITNYHYLANAVIGWQAGSLVYVNMQIGYATPILILSLC
LAQLSLVHLVLSLSESQAKALGQKTNLISAVFMILVILSSAAVIAIGSISFGLVI
PHLMKHETPHRYRLPLCAVSGASFMWVDIACRNLPPEYETPLGALVSLIGPCEL
WLIRRGRY"

```

```

gene
8188..9225
/gene="Sps1579"
8188..9225
/gene="Sps1579"
/notes="similar to GB:AAK33424.1 (AE006501) percent
identity 98 in 345 aa"
/codon_start=1
/evidence-not_experimental
/transl_table=11
/product="putative ferrichrome ABC transporter (permease)"
/protein_id="BAC64674.1"
/db_xref="GI:28811745"
/translation="MLFVWKKRRLLKSHRLLGLYTSLLILLVSLMGLSLSGESH
LSPDLIVPFGKSSHAISFIYINIRLPRILAAAGLSLALSGLLQRLTRNPLADS
GVLTIGTICAGLSLAIVVSPFPEQAHSHVLPFLFAMLANVVTTSYVWTLSPKQCID
PRLLITGVATYTHLSLVALVGHINRYKVDLVINMLSGQIGDDWPFILSVIAPELL
CWLUTYQAHFUNGIMLADNTAIGLGLPLNRRRLILVLAAGLGSALLVGNISFI
GLIAGHFSYLVGSNNHKLTIPIISILIGMILLVADTVGRVTVNGSNTQITGLVSLIGA
PYFLYIMAKTK"
complement(9270..9923)
/gene="Sps1580"
complement(9270..9923)
/gene="Sps1580"
/notes="similar to GB:AAL97173.1 (AE009984) percent
identity 98 in 217 aa"
/codon_start=1
/evidence-not_experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC64675.1"

```

Alignment Scores:

```

Pred. No.: 3 45e-129 Length: 323825
Score: 1585.50 Matches: 298
Percent Similarity: 85.13% Conservative: 53
Best Local Similarity: 72.33% Mismatches: 60
Query Match: 73.92% Indels: 1
DB: 1 Gaps: 1

```

US-09-469-200D-2 (1-417) x AP005146 (1-323825)

```

Qy 5 LysAsnLeuThrValValAlaPheSerIlePheThrValLeuLeuIleTyrValAsn 24
Db 303845 AAAAAAACHTTTAAATGTTTATCTTTATTTGTTGATATCTATCTGATTTATCTAAAT 303904
Qy 25 ValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeuLeuIleAlaTyr 44
Db 303905 ATGTATCTATTGGACACA---TCACTGTAGCAATTTATGGAGTAATATTAATACCTAT 303961
Qy 45 LeuLeuValLysMetSerLeuSerPhePheTyrLysProPheLysGlyArgAlaGlyGln 64
Db 303962 CTAGTTTAAACTTGGGATTATCTTCTTATGAGCATTTAAAGAAATCCACATGAC 304021
Qy 65 TyrLysValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSerLeuLeuGluThr 84
Db 304022 TATAAGTTGCTGCTGTAATCTCTTATAATGAAGTCCGAGTCATTATTAGAAACA 304081
Qy 85 LeuLysSerValGlnGlnInThrTyrProLeuAlaGluIleTyrValValAspGly 104
Db 304082 CTTAAAGTGTCTAGCAGACACCTATCCGTTATCAGAAATTTATATTGTTGATGGG 304141
Qy 105 SerAlaAspGluThrGlyLysArgIleGluAspTyrValArgAspThrGlyAspLeu 124
Db 304142 AGTTCAACACAGATGCAATACAAATTAATTGAAGAGTATGTAATAAGAGAAAGTGGATATT 304201
Qy 125 SerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArgHisAlaGlnAla 144
Db 304202 TGTCGAAGCATTTATGTTTACCGCTCCCTTGTCTATCAAGAAACGCTGCTCAAGCG 304261
Qy 145 TrpAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSerAspThrTyrIle 164
Db 304262 TGGCCATTTGAAGATCTGACGCTGACGTTTTTTTACCGTAGATTCAGATCTATATTC 304321

```


QY 165 TyrProAspAlaLeuGluLeuLeuLysThrPheAsnAspProThrValPheAlaLeu 184
 DB 304322 TATCAATGCCCTTGAAGAACTCTAAAGGCTTCAATGATGAGACGTTTATCTGCA 304381
 QY 185 ThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArgLeuThrAspIle 204
 DB 304382 ACAGGACATTGATGCTAGAACAGACAACTATCTATTAAAGCGCTTACAGATATC 304441
 QY 205 ArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThrGlyAsnIleLeu 224
 DB 304442 CGTTACGATAATGCCCTTGGGTGAGCGTCTGCTCAATCATTAACAGGTAATATTTA 304501
 QY 225 ValCysSerGlyProLeuSerValTyrArgArgGluValValProAsnIleAspArg 244
 DB 304502 GTTGTGCTACGACCATGAGTATTATTCGACGTGAGGATGATATTCCTCACTTACAGCGC 304561
 QY 245 TyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAspArgCysLeuThr 264
 DB 304562 TATAAAATCAACATCTCTAGGTTACCTGTTAGCATGGGATGATGATGTTTAA 304621
 QY 265 AsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLysCysIleThrAsp 284
 DB 304622 AATTATGCTATTGATTAGACGCACTGCTACCAATCAACAGCTAGATGATGATGAT 304681
 QY 285 ValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrpAsnLysSerPhe 304
 DB 304682 GTACCTTTCCAAATTAAGAGTTATTAAAGCAACAAATCGATGAATAAATCTTTT 304741
 QY 305 ArgGluSerIleLeuSerValLysLysIleMetAsnAsnPropheValAlaLeuThrPhe 324
 DB 304742 AGAGAATCTATTATCTCTCAAAAAATCTTCTTAATCCCATGCTGCTTATGACT 304801
 QY 325 IleLeuGluValSerMetPheMetLeuValTyrSerValValAspPheValGly 344
 DB 304802 ATTTCCGAATCGTATGTTATGATGTTGATGTTGCGCAATGGGAATCTTTGTTTAT 304861
 QY 345 AsnValArgGluPheAspTrpLeuArgValLeuAlaPheLeuValIleLeuPheVal 364
 DB 304862 CAAGCTATTCATAGACCTTATTAACTTTTGCTTTTATCCATCATCTTATCGTT 304921
 QY 365 AlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPheLeuLeuSerPro 384
 DB 304922 GCTTTATGCTGTAATGTTATGATATGATGCTCAACATCTGCTAGTTTATGTTATCCT 304981
 QY 385 PheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyrSerLeuPheThr 404
 DB 304982 CTGTATGGAATATTACACTGTTGTTGTTACGCGCCCTAAACATTTATCTTTATGCACC 305041
 QY 405 IleArgAsnAlaAspTrpGlyThrArgLysLysLeu 416
 DB 305042 ATTAAAAATACGAATGGGAACACGTAAGAGTC 305077
 RESULT 13
 AF065404
 LOCUS
 DEFINITION Bacillus anthracis virulence plasmid pX01, complete sequence.
 ACCESSION AF065404
 VERSION AF065404.1 GI:4894216
 KEYWORDS
 SOURCE Bacillus anthracis
 ORGANISM Bacillus anthracis
 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus
 cereus group.
 REFERENCE
 1 (bases 1 to 181654)
 AUTHORS Okinaka,R.T., Cloud,K., Hampton,O., Hoffmaster,A.R., Hill,K.K.,
 Keim,P., Koehler,T.M., Lamke,G., Kumano,S., Mahillon,J., Manter,D.,
 Martinez,Y., Ricke,D., Svensson,R. and Jackson,P.J.
 TITLE Sequence and organization of pX01, the large Bacillus anthracis
 plasmid harboring the anthrax toxin genes
 JOURNAL J. Bacteriol. 181 (20), 6509-6515 (1999)
 MEDLINE 99445483
 PUBMED 10515943
 REFERENCE
 2 (bases 1 to 181654)

AUTHORS Okinaka,R.T., Cloud,K., Hampton,O., Hoffmaster,A., Hill,K., Keim,P.,
 Koehler,T., Kumano,S., Lamke,G., Manter,D., Martinez,Y.,
 Ricke,D.O., Svensson,R. and Jackson,P.J.
 TITLE Direct Submission
 JOURNAL Submitted (14-MAY-1998) Life Sciences Division, Los Alamos National
 Laboratory, P443, LS-6, HRL-1, MS M888, Los Alamos, NM 87545, USA
 FEATURES
 Location/Qualifiers
 source
 1..181654
 /organism="Bacillus anthracis"
 /mol_type="other DNA"
 /strain="Sterne"
 /db_xref="taxon:1392"
 /plasmid="virulence plasmid pX01"
 complement(1127..1630)
 /note="Hypothetical protein, Methanobacterium
 thermoautotrophicum, (AE000831)"
 /codon_start=1
 /transl_table=11
 /product="pX01-01"
 /protein_id="AAD32305.1"
 /db_xref="GI:4894217"
 /translation="MEVLIPELILIAVLIPNSVKKHVPKWKAGKELIYRMLSKL
 DKSYVLIHNVITVYTGDTQIDHIVIAETGVVFTKNGYEWIGSEKAAATVQGI
 FRKSFQNFPHONYKHIAEWLIEQQLFCISMAFPHKGLARVNVHSEKHKYKL
 HNKFAQ"
 complement(1639..1644)
 RBS 2751..2755
 RBS 2760..3338
 CDS
 /codon_start=1
 /transl_table=11
 /product="pX01-02"
 /protein_id="AAD32306.1"
 /db_xref="GI:4894218"
 /translation="MLEHTLPSLVQLVPLIVSTGIYILLETDTENKWTESYEGIF
 AILITRFQVQNKSLISLQHEKFWKSLNIFLVSILVLFASIFTLKALFILOM
 TWILEYLCYIDRLPIHNLRYKVEICDKKAKIERIPNNEEHGTHITKTRDEQ
 YFLYAVYLSGLISPSIEEFVEYERNLMGR"
 complement(3389..4024)
 CDS
 /codon_start=1
 /transl_table=11
 /product="pX01-03"
 /protein_id="AAD32307.1"
 /db_xref="GI:4894219"
 /translation="MKREKMGIVLCTLVMLILLVPLQAQASVYDFNKLKGDVNS
 SPKGLANRKTTPKDDSPGSSIIQKTEGLYAPAFKGVNFISSIFVLGIIVMVAL
 IFKQNMKWATGSMFSPISMSNRLSVFLPSDFSENNITDFLDLFAKALIF
 TFMLLAGLRRLRHSTKQPSYRYWANRITGSCILCLMLAVLAPWLFANN"
 complement(4024..4653)
 CDS
 /codon_start=1
 /transl_table=11
 /product="pX01-04"
 /protein_id="AAD32308.1"
 /db_xref="GI:4894220"
 /translation="MNINSNNVKELEQTVRYKKQVQLRVSKSSNIHAYMIFP
 EFSEVKNTLEDYIENLSTVKYIDRKLSQOEIEKQNHSEFIMTCEVEDGVIDDE
 THFTVTKPTTANKFKVLCGLNGLSLKELHITVNNNINRIEIFYENETY
 VAYKEITANIMGAYTKANIYAADLYAAEGNEIKLKSIPKLAGDV"
 complement(4034..4039)
 RBS complement(4660..4665)
 RBS complement(4989..5309)
 CDS
 /codon_start=1
 /transl_table=11
 /product="pX01-05"
 /protein_id="AAD32309.1"
 /db_xref="GI:4894221"
 /translation="MELETEDMYQQYVSIPPETPVFPQQRVADVKMDQFVWRT
 TKYEOGQLFRVIMNSNNRWIHTGEILPENQLLYLKEVDSYSLTVLEWSNRQ
 LGNI"
 complement(5318..5323)
 RBS complement(5720..6220)
 CDS
 /codon_start=1
 /transl_table=11
 /product="pX01-06"

```

/protein_id="RAD32310.1"
/db_xref="GI:4894222"
/translation="MLGDTGLQTLIIQNIKSNQNLKGSTEEWTKAASGESKRVLANI
GNVITVFESIQITWNVEKLPFGTMDVAFMKEAQFYQKAVECIKKEQVPGVRVQKNV
LEVYDGEIEDIRALTTLNTDKRLVQKQDITFWSVIVPEVDPSSKKKGKQKNIIDQY
EFDLFG"
RBS
CDS
complement(6231..6236)
complement(6544..8352)
/notes="reverse transcriptase, IS629, E. coli, (AB011549);
reverse transcriptase, IS629 (574 aa) E. coli (AB011549);
306/543 positive aa (56%)"
/codon_start=1
/transl_table=11
/product="pX01-07"
/protein_id="RAD32311.1"
/db_xref="GI:4894223"
/translation="MNTVMKSAVSQVTVQNSIQWKEIENYVVKLRRIYRAQFGN
KRVKRLQMLRSKANLLISIRVQINQKRTAGDITTPEDRVKFLHLLKG
SVRIKAFVPRKATIPKNGKGRPLGIPVKIDRFQNNVKNALPQWECREPSYGF
RPRKSAHADAMALFKLSRGTNRAMIFGDFGQCFDNLNHEHILSCIEGFPYSNAIQ
WLAGCTIDNKTYKTECTPGGIIISPLANIALHGMKELGVRHFKRQGLMYPD.
SIGIVRYADDFVYVCSKEAESWYAKLQPYLDKRGKLAELAEKTRVHHTQDFDIFG
NFRQYFKGSGDLFKPSNSQSKAKKEISELPSKHSRSGIGQLIRKLNPIVITGIAN
WSPVAKVIYGDIDSYVHKVMHHLKXHHKRGARWINKKYFHPDHTGVSDKRLITD
PDNHKQLIRMTPIVIRHLIKYKNSPDANLKKYFAKDEKIFNRFTNSKRLAK
KTKYKRCVNSLVGEPELESNHIVPKVTGGKDEYDNLLELHCHSKQHALLLEWGN
GKOLPKVQALYSOKIPINSKAVRMLNTFKFKY"
RBS
CDS
complement(8362..8367)
complement(8956..11541)
/codon_start=1
/transl_table=11
/product="pX01-08"
/protein_id="RAD32312.1"
/db_xref="GI:4894224"
/translation="MLQELIKQDTPQEQTPALQLLETGTASHSAPCFAMAINHHREI
QFAYLGANDSTLKSRAISWGTSLRYFEGEKRESHYTLDKVNVISGQFPEINTQ
TINKKALIAFSKQLETKIYVAIDEAPENQVDFLMAPPYGLILEEWAKEPIEEMIT
KNLLQPLEYFQNFESLSIAQVILKEDCKEFLSEMSIRSKQCFQPGGTGEKINKI
QDLNLELYFSEVMDLVKIDLEPLQHPMKQEILSHEDTYKRLFLTPVAHATGAAGA
LQVQKGLIIQGGMSKSSIMATVDGPHLTGKGYRTCVFPPTLTAKWAKEIRH
LIPADVHILKRTEDLIRHQSLQTPGKPEKTEFFVISFTIMRGDSIKOMPIYKO
IALSKSEEVORYKNGYCPDGGAKLEKTSLSIMVHQANGEQREVCQYKDFGSL
DSKTNKNSVACNSNIWSVKTYSKFWKHTFATKCNKLVQALKEGNKPLQKLELEN
RVYDPAKQSGRAYKATVETIRKMKHFDALICDEVHELKETAQGTALGSLAKV
CKYVAGTGLTFGRASDVFTLWRLFPQKVKSGEYSFSEWNETNGTERTYHYD
GEVNTISASGAGQTLDTKVPVPGISPYVTOFLMDTINVRLKQWPNVPLVNPY
ILDVMSQKQAYEHMKSEPEKLEAAGKTDPAGKMLSVIRGTGNAYPDNMHPNY
LDGAKELVNSDKFKEDFCIQPEKELQSLITELSEGRITVYVCDTGSTVKRED
VQPLQKWEQIPGAKVATLRTNVTVPKRSWLKEQVNSGVNVIICSLVLYGLDC
AIRSR"
RBS
CDS
complement(11549..11554)
complement(11613..11840)
/codon_start=1
/transl_table=11
/product="pX01-09"
/protein_id="RAD32313.1"
/db_xref="GI:4894225"
/translation="NSVLLGIGSLIIWVEFSNRPRIATVLTVICASQLGAGTHLE
VLEQFVYFGVGDMLSAEQLSPENKRFIDF"
RBS
CDS
complement(11847..11852)
complement(11856..12947)
/codon_start=1
/transl_table=11
/product="pX01-10"
/protein_id="RAD32314.1"
/db_xref="GI:4894226"
/translation="MRLGNDLAAGFYPTPLTEGKILYVLLQMESERAYACDPCCGSG
TILKSFADETKQGVQFHYTGVDLNRNITGQAKQLDVVYKSSFESMISHNYPFLIF
LNPNYTELETKENAKSEKMFNFKRAHLYLQEGIMVYIPYDFRARDQISYLAKN
YEETGLMRFDEENEEFQKQCFVIGKRAATKNDTFNDRFANFCENMSDLDVRKH
VSTLAQMVNHKWTIPELRHQNKLIPTSRVDYKDAYGKIGKSGIASLKKRLNGY
LLOGEKTAASRAKMPIASSQGLGLLITGVADGLLEGDTLHVRGSENVYEHSEFQ
EKTKEVVLQKRAKEITATPTGVKELY"

```

```

RBS
CDS
complement(12956..12961)
complement(13209..13562)
/codon_start=1
/transl_table=11
/product="pX01-11"
/protein_id="AAD32315.1"
/db_xref="GI:4894227"
/translation="MEDNOKSYAAPVMSSESRAVYNMRGCFEAILIPDGKFOYFKP
HALNKKRATQALYATIEGQEMKGVYILWRFVGGQSFYSGTKIYAPKRLILRLK
DIFFDYEVYEPKK"
RBS
RBS
CDS
complement(13574..13579)
14226..14231
14237..14929
/codon_start=1
/transl_table=11
/product="pX01-12"
/protein_id="AAD32316.1"
/db_xref="GI:4894228"
/translation="MLKNDILCSFTVFSLIFFVPTQDITMTKVIYFVINGLGISI
PIVAMRMOLKOTIKLDFDKEMKEITTKKEITNFVGLLKQYIQGDKLRSEHYIN
RESEMGILYREFDWDKAYAIIPINWAKLIDDYASILNLEELDPIHQIVLSY
VTFACEETLNAIFSLGSGSEKDEKVKLLSRNIQKNIVLSAGFEKVFKEVEETITE
FTFVLVQNKRG"
CDS
complement(15040..19002)
/notes="erythrocyte invasion and possible binding protein;
Shptry protein (2401 aa), plasmodium yoelii (U36927);
520/1194 positive aa (43%)"
/codon_start=1
/transl_table=11
/product="pX01-13"
/protein_id="AAD32317.1"
/db_xref="GI:4894229"
/translation="MAKKEIPDIYKYEDIKAAKGNLIIYFVGAGISMLGLPSWK
YAQELIDAEKGVDSLEYKLEEDPKILTKMLFDEKEIIPKSNALFEKPKN
EYKAIYKLYSNAIYITNYDECLDVLITPQSDNEKYSMDLSEKNLEQSET

```

Alignment Scores:

```

Pred. No.:      8.53e-82      Length:      181654
Score:          1052.50      Matches:      213
Percent Similarity: 66.28%      Conservative: 74
Best Local Similarity: 49.19%      Mismatches: 121
Query Match:      49.07%      Indels:      26
DB:              1          Gaps:      5

```

US-09-469-200D-2 (1-417) x AF065404 (1-181654)

```

QY 3 ThrLeuLysAsnLeuLeuThrValValAlaLapheSerilePheTrpValLeuLeuTyr 22
Db 111183 TCITTAATAAGCCCTTCATATTACTTGTGTTTCTTTATGATTAGTAAGTACTAT 111242

QY 23 ValasnValTyrLeuPheGlyAlaLysGlySer -----LeuSerileTyrGly 38
Db 111243 ATTAAC -----TTTACCATTCTGTGGAAGTTTATTGTTTATTAGGCATATATGA 111293

QY 39 PheLeuLeuLeuAlaTyrLeuLeuValLysMetSerLeuSerPhePheTyrLysProPhe 58
Db 111294 ACGTGATGCGATTACTTACTTGGAAAAACAATCATATCATTT-TTTTATCGCAAAAT 111352

QY 59 LysGly---ArgAlaGlyGlnTyrLysValAlaLallelleProSerTyrAsnGluAsp 77
Db 111353 ACAGTGATGAAGTCCCAATATGAAGTGGCGAGTAGTGGTCCCTCATATATGAAGT 111412

QY 78 AlaGluSerLeuGluThrLeuLysSerValGlnGlnGlnThrTyrProLeuAlaGlu 97
Db 111413 GCAAGTCATTGTTAATAACAATAATATAGCGTTTTAGCTCAAGATTATCCAATTCATGAA 111472

QY 98 IleTyrValValAspAspGlySerAlaAsp-----GluThrGlyIleLysArg 113
Db 111473 ATTTCTTTGTCATGATGCTAGTAAGATAAATCGGCTTATCAAGTAGCAGCTTAAATG 111532

QY 114 IleGluAspTyrValArgGlyThrGlyAspLeuSerSer----- 126
Db 111533 AGGGAGGAGACTTCTTAGAAGTCAACGAGAAATTCGCTACAACTAAGAAATATTGTTCT 111592

```


CDS
/gene="BXA0005"
2291..2875

/gene="BXA0005"
/note="identified by Glimmer2; putative"

/codon_start=1
/transl_table=11

/product="hypothetical protein"
/protein_id="AAM25962.1"

/db_xref="GI:20520080"

/translation="MLKLVNNTGTPEGDSAFNYETCKDCVYLSKSDSCSLQAA
DSNNPLVECCNFEYEYETEDPLNSHQREKAQILAEAPAPQSDIPHD
AIWTSVNGCWLINDSKKRALAEQVPGDNNSKKYASPPPLHDGAKGIRHYM
CWYNEKGIGCEYHLLIGVTMTAEGYKGLGPNY"

gene
2895..3473

CDS
/gene="BXA0006"
2895..3473

/gene="BXA0006"
/note="identified by Glimmer2; putative"

/codon_start=1
/transl_table=11

/product="hypothetical protein, (px01-02)"
/protein_id="AAM25963.1"

/db_xref="GI:20520081"

/translation="MLEHTILSVQLVPLVSGLIYILLETETKWTWFSVEGIF
ALIPTRFOVFNKKSIIINQHKWFKUNIFLVISFLVLFASPTLLKAFIOM
TVILFLEYLGLIDRLPHNLXRVKEIIDCKKQKLTIERIPNNEHGHTIKTRDQ
YFLTYLVSGLSPISSEFEYERNLMGR"

gene
complement(3524..4111)

CDS
/gene="BXA0007"
complement(3524..4111)

/note="identified by Glimmer2; putative"

/codon_start=1
/transl_table=11

/product="hypothetical protein, (px01-03)"
/protein_id="AAM25964.1"

/db_xref="GI:20520082"

/translation="MILLVPLQAQASVVDNFKLGNVSSFKDGLNRTYTKK
DMSGSIQKIEGLYAFKQVNIIFSIVGLVYVVMALITKNGQMKWATGSMI
FSFISMSMRLSYFLSDVSFNFIDFLQFKALIFTPEMLIAGIRLRRIHE
STKQPYRWANRVITGSLCLMLAVLPWFANN"

gene
complement(4159..4788)

CDS
/gene="BXA0008"
complement(4159..4788)

/note="identified by Glimmer2; putative"

/codon_start=1
/transl_table=11

/product="hypothetical protein, (px01-04)"
/protein_id="AAM25965.1"

/db_xref="GI:20520083"

/translation="WNINSNNVVEKLSQTVRYEKKQVQLRVKSSNTHAYNIFP
EFSVKISNTEDEYIELNSTSVKVIDRKLQQEIKQKHSFIMTCVEEDGVIDE
TIFHTYKPTKPTKAVKLVGLGNRLDSKLHETVLNNINRIETEFYENETRY
VAVEKEEIANAGYATKANIYAAADLYAASGNEIKLSIKPKLAGDV"

gene
complement(5124..5837)

CDS
/gene="BXA0009"
complement(5124..5837)

/note="identified by Glimmer2; putative"

/codon_start=1
/transl_table=11

/product="hypothetical protein, (px01-05)"
/protein_id="AAM25966.1"

/db_xref="GI:20520084"

/translation="MFTKIFIGLFCYFVIVGNSKKFLREAAKKEMENKPLRPVY
WCPRENKIQEORITVPTKEIDKVDVATSVHKNLGEVSDRLKHEIDYE
EVEVAFDHTPTPRAYDIDHPDSEDIPVELLEYDMYQQEYVSIPEPIPPVQ
QRVADVDMDFVWTKVLEQEGQLRVINMNNKRWIHTTEGELLPENQLIMKLEV
DSPTSYTLVWNSRQLGNI"

gene
complement(5854..6561)

CDS
/gene="BXA0010"
complement(5854..6561)

/gene="BXA0010"
/note="identified by Glimmer2; putative"

/codon_start=1
/transl_table=11

/product="hypothetical protein, (px01-06)"
/protein_id="AAM25967.1"

/db_xref="GI:20520085"

/translation="MIFYOLPYSFALNOAARHCRIGQTKOCTFVLGYRETFOOM
AOLIAQKNAEAMNGEATISGLNMLGDTGLDQLILLQIKSGNQLKSTEWTKAA
SESKRVLANIGNVTTPVEESIEQFTNWVKLPQMDVAFMKEAQFQKAVECIKKE
QVPGYRVQNVLEVDEPIEDIIIRALTTLNTRKMLVQKODITFDWDSVIVEVPSKKKG
KQKNKIDGGYEFDLFG"

gene
complement(6678..8486)

CDS
/gene="BXA0011"
complement(6678..8486)

/note="identified by match to PFAM protein family HMM
PF00078"

/codon_start=1
/transl_table=11

/product="reverse transcriptase/endonuclease protein,
(px01-07)"

/protein_id="AAM25968.1"

/db_xref="GI:20520086"

/translation="MNTVMKSAVSPQVQWNSIQWKEIYVKKLRQIRYAEQFGN
KRVKRLQRLAKKANLLLSIKRVQINQKKRTAGIDGITTNPEDRVKLFHLKYG
SVRNTKAPVYKRAVIPKNGKKRPLGIPVIGKRIKQVKNALPEQWECRFESMSYGF
RPKSAHDAMANFLKLSRGTNRWIFEGDFQCFDNLNHHILSCIEGFPYSNAINQ
WLNAGCIDNKTFYKTETGPGGIIISPLANIALHGMKEKELGVYHPRDGMALYDP
SIGTVYADDFVIVCSKEEASMTALQYLDKGLKLAETRVVHITDGFDFLPG
NRPQYTKESQDFIKPSNQSKAKKEKISEIFKSHRSGISQGLIRKLNPTVITGANY
WSPVAVYIGDIDSVYHKVNMHLKYKHKRGARWINKYFHPDHTGYSDQKLLTD
PDNKNQRLRMWTPVIRVHLIKYKNSPDDANLKYFAKDEKIFNFRNTNSKRLAK
KTKYKRNNSLVGEESLNIHVPKIVGKDEYDNLLELLHCSCHQHALLWVGN
GKOLPKVOAYLKSQKIPINSKAVRTMLNFKKPY"

gene
complement(8991..9083)

CDS
/gene="BXA0012"
complement(8991..9083)

/note="identified by Glimmer2; putative"

/codon_start=1
/transl_table=11

/product="hypothetical protein"
/protein_id="AAM25969.1"

Alignment Scores:

Pred. No.:	8.53e-82	Length:	181677
Score:	1052.50	Matches:	213
Percent Similarity:	66.28%	Conservative:	74
Best Local Similarity:	49.13%	Mismatches:	121
Query Match:	49.07%	Indels:	26
DB:	1	Gaps:	5

US-09-469-200D-2 (1-417) x ABO11190 (1-181677)

Qy 3 ThrLeuLysAsnLeuLeuThrValAlaPheSerIlePheTrpValLeuLeuLeuTyr 22

Db 111234 TCITTAATAATGCCCTTCATATTACTTGTGTTTTTCTTTATTGTTAGTACCAATGCTACTAT 111293

Qy 23 ValAsnValTyrLeuPheGlyAlaLysGlySer-----LeuSerIleTyrGly 38

Db 111294 ATTAAC-----TTTACCATTCTGGAGAGTTTGTGTTGTTTGGGATATATGGA 111344

Qy 39 PheLeuLeuLeuAlaTyrLeuLeuValLysMetSerLeuSerPhePheTyrLysProPhe 58

Db 111345 ACAGTGATGTGATTACTTACTTACTTGGAAAAACAATCATATTATTTATCGACCAATT 111403

Qy 59 LysGly---ArgAlaGlyGlnTyrLysValAlaAlaIleLeuProSerTyrAsnGluAsp 77

Db 111404 ACAGGTGATAAGTCCCAATATGAAAGTGCAGTAGTGGTCCCTCATATATGAAAGT 111463

Qy 78 AlaGluSerLeuGluThrLeuLysSerValGlnGlnThrThrProLeuAlaGlu 97

```

Db      111464  GCAGTCTATTGTTAATACAAATTAAGCGTTTAGCTCAGATTATCCAAATTCATGAA 111523
QY      98      IleTyrValValAspAspGlySerAlaAsp-----GluThrGlyIleLysArg 113
        |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      111524  ATTTCCTTGTGATGATGCTAGTAAGGATAAATCGCTTATGAAGTAGCACTTAAATG 111583
QY      114      IleGluAspTyrValArgAspThrGlyAspLeuSerSer----- 126
        |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      111584  AGGGAGGAACCTCTTCACTCACTCAAGCAAGAAATTCGCTGACAACTAAGAAATATTTGTTCT 111643
QY      127      -----AsnValIleValHisArgSerGluLysAsnGlnGlyLysArg 140
        |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      111644  GAAATATTAGTATTCTCTGACTTATCGTACATCGTTTACCTAAGAAATTCGCGGAAAGA 111703
QY      141      HisAlaGlnAlaTrpAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSer 160
        |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      111704  CATGCTCAATATGGCTTTTAAACGGCAACACAGACAGATGCTATTGTACCATTTGATTCGA 111763
QY      161      AspThrTyrIleTyrProAspAlaLeuGluGluLeuLysThrPheAsnAspProThr 180
        |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      111764  GATGCTGATTGTTCCTCCAAATGCTGTAGAGAGTTATTGAAACCTTTAATGATGAAAAA 111823
QY      181      ValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArg 200
        |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      111824  GTAATGGCCACAACTGGTCACGTGAACATTCGTAATAGAAATGATAATTTATTAACAAAA 111883
QY      201      LeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThr 220
        |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      111884  CTAATTGATATGGCTTATGACAATGCGTCCGTGGAGCGGCGACAGTCCGTAACA 111943
QY      221      GlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgGluValValPro 240
        |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      111944  GGAATGTTCTTGTGTAGTGGCGCTTAAGTTGTTATTCGTAGAGAACTAATACTGAA 112003
QY      241      AsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAsp 260
        |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      112004  AATTTAGAACAATTATGGAAGTCAGATGTCCTTGGTGGAGAGGTGCAGTTGGAGATGAT 112063
QY      261      ArgCysLeuThrAsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLys 280
        |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      112064  AGATGCTCAACTAATATGCTATTGTTGAAGGGGAAAAACAGTTTATCAATCCACTGCTGGA 112123
QY      281      CysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTTrpAsn 300
        |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      112124  TGTATTACTGATGCTCCACTACATTAAACAATTCCTTAACACGAACTACCTTGGAAAC 112183
QY      301      LysSerPhePheArgGlnSerIleLeuSerValLysLysIleMetAsnAsnProPheVal 320
        |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      112184  AAGTCATTTTTAGAGAAAGTTTAAATTTCACTTGGCATTTGGTATGAAAAACCAAAATGTT 112243
QY      321      AlaLeuTyrThrIleLeuGluValSerMetPheMetMetLeuValTyrSerValAsp 340
        |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      112244  CTTGTTTGGCAATTTTCGAATATCGTTATGGATTTTATTTGGCGTTTCCCTACTTCTA 112303
QY      341      PhePheValGlyAsnValArgGluPheAspTrpLeuArgValLeuAlaPheValIle 360
        |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      112304  AGTATTATTCTCAAGGCAAGTCATGCTAGGTTAATTTGGTGGCTTTATTATTGGGTTAT 112363
QY      361      IlePheIleValAlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPhe 380
        |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      112364  ATTTCATTAGCTGATATGCTAGAATGATTTTATCTATTATTAACATCCCTCTTTTC 112423
QY      381      LeuLeuSerProPheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyr 400
        |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      112424  TTACTGGCGCCATTATATGGAATCTCCATGATTTACACACTATTACCTATACGCTTTAT 112483
QY      401      SerLeuPheThrIleArgAsnAlaAspTrpGlyThrArg 413
        |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
Db      112484  GCTTTACTAACTATTAAATCTAATGTTGGGGGAACAGT 112522
QY      15
AF518732
LOCUS      AF518732      1893 bp      DNA      linear      SYN 28-OCT-2002

```

```

DEFINITION Synthetic construct has operon deletion mutant, partial sequence.
ACCESSION AF518732
VERSION AF518732.1 GI:24415825
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 1893)
AUTHORS Walker, J.A. and Timoney, J.F.
TITLE Construction of a stable non-mucoid deletion mutant of the
JOURNAL Vet. Microbiol. 89 (4), 311-321 (2002)
MEDLINE 2272956
PUBMED 12383640
REFERENCE 2 (bases 1 to 1893)
AUTHORS Walker, J.A. and Timoney, J.F.
TITLE Direct Submission
JOURNAL Submitted (05-JUN-2002) Veterinary Science, University of Kentucky,
108 Gluck Equine Research Center, Lexington, KY 40546, USA
FEATURES
source
1..1893
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
<1..>1893
/gene="has operon deletion mutant"
/note="derived from Streptococcus equi Pinnacle strain"
<1..>423
/gene="hasA"
/note="hyaluronate synthase"
430..1137
/note="derived from vector pTW100"
<1138..1356
/gene="hasB"
/note="UDP-glucose dehydrogenase"
1406..>1893
/gene="hasC"
1406..>1893
/gene="hasC"
/codon_start=1
/transl_table=11
/product="UDP-glucose pyrophosphorylase"
/protein_id="AAN59302.1"
/db_xref="GI:24415826"
/translation="MPKVRKALIPAAIGLGRFLPATKALAKEMLPIDVKPTIQIFVEE
AKGLAEILIVTCKAKRSIEDPDSNELEYLNQAKGTLLKLVDETAINLAEIR
CSHPGLGDVLAQKAFVGNPEFVVLGDDLDITNPSAKPLTKQLIEDICTHASTI
AV"
BASE COUNT 559 a 427 c 385 g 522 t
ORIGIN
Alignment Scores:
Pred. No.: 4, 2e-57 Length: 1893
Score: 745.00 Matches: 147
Percent Similarity: 88.95% Conservative: 6
Best Local Similarity: 85.47% Mismatches: 19
Query Match: 34.73% Indels: 0
DB: 12 Gaps: 0
US-09-469-200D-2 (1-417) x AF518732 (1-1893)
QY 159 AspSerAspThrTyrIleTyrProAspAlaLeuGluLeuLysThrPheAsnAsp 178
Db 1 GACTCAGATACTTATATCTACCTCGTGTAGAGGAGCTGTTAAAGACCTTTAATGAC 60
QY 179 ProThrValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeu 198
Db 61 CCAACTGTTTTTCTGCGACGGGTCACCTTAATGTGAGAAATAGACAAACCAATCTCTTA 120
QY 199 ThrArgLeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSer 218
Db 121 ACACGCTTGACAGATATTGCTGATGATATGCTTTTGGCGTTGAGCAGCTGCCCATCA 180

```

QY 219 ValThrGlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgGluValVal 238
 DB 181 GTTACGGGTAAATATCTGTTGCTCAGGCCACCTAGCGTTACAGCCGAGGTGTT 240
 QY 239 ValProAsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGly 258
 DB 241 GTTCCTAATATAGACAGATACATCAACACAGACCTCCCTGGGTATTCCTGTAAGTATCGGT 300
 QY 259 AspAspArgCysLeuThrAsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThr 278
 DB 301 GATGACAGGTGCTTGACCAACTATGCAACTGATTTAGGAAAGACTGTTTATCAATCCACT 360
 QY 279 AlaLysCysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArg 298
 DB 361 GCTAAAGTATTACAGATGTTCTTGACAGATGCCCTACTTACTTGAGCAGCAAAACCGC 420
 QY 299 TrpAsnLysSerPhePheArgGluSerIleIleSerValLysLysIleMetAsnAsnPro 318
 DB 421 TGGGATCCACTAGTTCTAGAGCGCGCCACCGCGGTGGAGCTCCAGCTTTTGTTCCT 480
 QY 319 PheValAlaLeuThrPheIleLeuGluValSerMet 330
 DB 481 TTAGTGGGTAAATTGCGCGCTTGGCGTAATCATG 516

Search completed: October 2, 2003, 14:19:11
 Job time : 5279 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2003, 09:01:58 ; Search time 4785 Seconds
(without alignments)
10721.144 Million cell updates/sec

Title: US-09-469-200D-1

Perfect score: 1254

Sequence: 1 atggaacattataaaacct.....cacgtataaaattattataa 1254

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl :

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rod.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1254	100.0	1254	1	AF023876 Streptococcus
2	1254	100.0	1254	6	BD087249 Hyaluronase
3	1254	100.0	1254	1	AY173078 Streptococcus
4	1228.4	98.0	7268	1	AF347022 Streptococcus
5	623.2	49.7	3466	1	SUB242946 Streptococcus
6	604.6	48.2	1350	1	STRHASAO
7	604.6	48.2	1440	1	STRHASA
8	604.6	48.2	1512	6	AR232201
9	604.6	48.2	10851	1	AE010123 Streptococcus
10	601.4	48.0	9844	1	AE006637 Streptococcus
11	598.2	47.7	53896	1	AE014172 Streptococcus
12	598.2	47.7	323825	1	AF005146 Streptococcus
13	409.2	32.6	1893	12	AF518732 Synthetic
14	282.8	22.6	181654	1	AF065404 Bacillus
15	282.8	22.6	181677	1	AE011190 Bacillus
16	99.4	7.9	6544	1	AF284858 Bradyrhizobium
17	73.2	5.8	6014	1	RGNOD
18	72.2	5.8	33038	1	AF222753 Bradyrhizobium
19	71.8	5.7	110000	1	REU80928.0 Bradyrhizobium
20	70.2	5.6	230573	1	AF322013 Bradyrhizobium
21	70.2	5.6	299600	1	AF005942 Bradyrhizobium
22	70	5.6	813	1	AF217269 Rhizobium
23	68.8	5.5	843	1	AF217262 Rhizobium
24	68.8	5.5	845	1	AF217263 Rhizobium
25	68.8	5.5	858	1	AF217268 Rhizobium
26	67.2	5.4	871	1	AF217265 Rhizobium
27	66.8	5.3	808	1	AF217264 Rhizobium
28	66.6	5.3	4317	1	ACNOD1
29	66.6	5.3	9549	1	RHANODABC
30	66.6	5.3	152050	1	ML0672113 Mesorhizobium
31	66.6	5.3	346820	1	AF003008 Mesorhizobium
32	65.8	5.2	933	1	AF217266 Rhizobium
33	65.4	5.2	825	1	AF166849 Rhizobium
34	65.4	5.2	839	1	AF166843 Rhizobium
35	65.4	5.2	839	1	AF166846 Rhizobium
36	65.4	5.2	839	1	AF166847 Rhizobium
37	65.4	5.2	900	1	AF166848 Rhizobium
38	65.4	5.2	3403	1	RGNODABCD
39	65.2	5.2	894	1	MCI457923
40	65.2	5.2	894	1	MME457924
41	65.2	5.2	894	1	MME457928
42	65.2	5.2	894	1	SSP457926
43	65.2	5.2	894	1	SSP457927
44	65.2	5.2	934	1	AF217271 Rhizobium
45	65	5.2	2627	1	RLNODABC Rhizobium

ALIGNMENTS

RESULT 1	AF023876	1254 bp	DNA	linear	BCT 15-OCT-2002
LOCUS	Streptococcus equisimilis hyaluronan synthase gene, complete cds.				
DEFINITION	Streptococcus equisimilis				
ACCESSION	AF023876				
VERSION	AF023876.1	GI:2655099			
KEYWORDS	Streptococcus dysgalactiae subsp. equisimilis				
SOURCE	Streptococcus dysgalactiae subsp. equisimilis				
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus				
REFERENCE	1 (bases 1 to 1254)				
AUTHORS	Kumari,K. and Weigel,P.H.				
TITLE	Molecular cloning, expression, and characterization of the authentic hyaluronan synthase from group C Streptococcus				

Pred. No. is the number of results predicted by chance to have a

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

equisimilis
J. Biol. Chem. 272 (51), 32539-32546 (1997)
98070433
9405467
2 (bases 1 to 1254)
Weigel, P.H. and Kumari, K.
Direct Submission
Submitted (10-SEP-1997) Biochemistry & Molecular Biology,
University of Oklahoma Health Science Center, P.O. Box 26901, BMSB
860, Oklahoma City, OK 73190, USA
location/Qualifiers
1. 1254
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="genomic DNA"
/sub_species="equisimilis"
/db_xref="taxon:119602"
1. 1254
/codon_start=1
/transl_table=11
/product="hyaluronan synthase"
/protein_id="AAB87874.1"
/db_xref="GI:2655100"
/translation="MRTLKMLTVAFSPFWLLIYVNVLEKAGSLIYGLLIAY
LVYMSFFRPFKRGAGQIKVAALIFSYNEDRSLELTKSVQVQYFPLAEIYVD
DGSADGSKRIEDYVNDTGILSSNVVHERSEKQGRHAQAWAFERSDADVELTVDS
DTYIYDALEELKTFNDVFAATGHLNVRQTNLLTLDIRYNAGFVERAQS
VTGNILVSGPLSVREVVPVNDIRYINOTFLGIPVSGIDDRCLTNATDLGKTVYQ
STAKCTIDVPKMSYLLKQONRNKSPFRESIISVKKIMNPFVALWTILEVSEFNL
VYSVDFVGNVREDFMLRVLAFLVLIIVALLCRNIHYMLKHLPLSFLSPYGVHLF
VLQPLKLSLFIIRNADWGRKKLL"

FEATURES
source

CDS

BASE COUNT
ORIGIN

Query Match 100.0%; Score 1254; DB 1; Length 1254;

Best Local Similarity 100.0%; Pred. No. 1.5e-308;
Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGACATTAATAAACCTCATACACTGTTGTGGGCTTTAGTATTTTGGGTACTGTTG 60
DB 1 ATGAGACATTAATAAACCTCATACACTGTTGTGGGCTTTAGTATTTTGGGTACTGTTG 60
QY 61 ATTACGTCATCTTATCTCTTTGTGCTAAAGNAGCTTGTCATTTATGCTTTTGG 120
DB 61 ATTACGTCATCTTATCTCTTTGTGCTAAAGNAGCTTGTCATTTATGCTTTTGG 120
QY 121 CAGATAGCTACCTATTAGTCAAAATGTCCTTATCTCTTTTACAAGCCATTAAAGGGA 180
DB 121 CAGATAGCTACCTATTAGTCAAAATGTCCTTATCTCTTTTACAAGCCATTAAAGGGA 180
QY 181 AGGCTGGGCAATATAGGTGCGAGCCATTATTCCTCTTATACGAAAGATCGAGTCA 240
DB 181 AGGCTGGGCAATATAGGTGCGAGCCATTATTCCTCTTATACGAAAGATCGAGTCA 240
QY 241 TTGCTAGAGACTTAAAAAGTGTTCAGCAGCAAACTATCCCTAGCAGAAATTTATGTT 300
DB 241 TTGCTAGAGACTTAAAAAGTGTTCAGCAGCAAACTATCCCTAGCAGAAATTTATGTT 300
QY 301 GTTGACAGTGAAGTCTGTGAGACACAGGTATTAAAGCGCATTAAGACTATGCGCTGAC 360
DB 301 GTTGACAGTGAAGTCTGTGAGACACAGGTATTAAAGCGCATTAAGACTATGCGCTGAC 360
QY 361 ACTGTGACCTATCAGCAATGTCATTGTCATCGGTGAGAGAAAATCAAGGAAGGCT 420
DB 361 ACTGTGACCTATCAGCAATGTCATTGTCATCGGTGAGAGAAAATCAAGGAAGGCT 420
QY 421 CATGACAGGCTGGGCTTTGAAAGATCAGACGCTGATGCTTTTGGACCGTTGACTCA 480
DB 421 CATGACAGGCTGGGCTTTGAAAGATCAGACGCTGATGCTTTTGGACCGTTGACTCA 480
QY 481 GATACCTATATACCTGATGCTTTTGGAGAGTGTGTTAAACCTTTATGACCCAACT 540
DB 481 GATACCTATATACCTGATGCTTTTGGAGAGTGTGTTAAACCTTTATGACCCAACT 540

QY 541 GTTTTGTGCTGCGAGGCTCACCTTAATGTCAGAAATAGACAAACCAATCTCTTAACAGC 600
DB 541 GTTTTGTGCTGCGAGGCTCACCTTAATGTCAGAAATAGACAAACCAATCTCTTAACAGC 600
QY 601 TTGACAGATATTCCTATGATATCTTTTGGGCTTGAACAGAGTGCCTTAAATCCGTTACA 660
DB 601 TTGACAGATATTCCTATGATATCTTTTGGGCTTGAACAGAGTGCCTTAAATCCGTTACA 660
QY 661 GGTAAATATCTTGTGCTCAGTCCGCTTGAAGCTTTTACAGACGCGAGGTGTTGTCCT 720
DB 661 GGTAAATATCTTGTGCTCAGTCCGCTTGAAGCTTTTACAGACGCGAGGTGTTGTCCT 720
QY 721 AACATAGATAGATACATCAACAGACCTTCTGCTGGTATTCTCTGAAGTATTGGTATGAC 780
DB 721 AACATAGATAGATACATCAACAGACCTTCTGCTGGTATTCTCTGAAGTATTGGTATGAC 780
QY 781 AGGTGCTTGACCAACTATGCAACTGATTAGGAAGACTGTTTATCAATCCACTGCTAAA 840
DB 781 AGGTGCTTGACCAACTATGCAACTGATTAGGAAGACTGTTTATCAATCCACTGCTAAA 840
QY 841 TGTATTACAGATGTTCTGCAAGATGCTACTTCTTGAAGCAGCAAAACCGCTGGAAC 900
DB 841 TGTATTACAGATGTTCTGCAAGATGCTACTTCTTGAAGCAGCAAAACCGCTGGAAC 900
QY 901 AAGTCTCTTTTAGAGAGTCCATTAATTTCTGTTAAGAAATCATGAACAATCCTTTGTA 960
DB 901 AAGTCTCTTTTAGAGAGTCCATTAATTTCTGTTAAGAAATCATGAACAATCCTTTGTA 960
QY 961 GCCTATGACCACTATGAGGTGCTATGTTATGATGCTGTTTATCTGTGTTGAT 1020
DB 961 GCCTATGACCACTATGAGGTGCTATGTTATGATGCTGTTTATCTGTGTTGAT 1020
QY 1021 TTCTTTTAGGCAATGTACAGAAATTTGATTTGGCTCPAGGCTTTTAGCCTTCTGTTGATT 1080
DB 1021 TTCTTTTAGGCAATGTACAGAAATTTGATTTGGCTCAGGCTTTTAGCCTTCTGTTGATT 1080
QY 1081 ATCTCATTTGTCCTGTGTCGGAACATTCATTACATGCTTAAGACCCGCTGTCCTTC 1140
DB 1081 ATCTCATTTGTCCTGTGTCGGAACATTCATTACATGCTTAAGACCCGCTGTCCTTC 1140
QY 1141 TTGTTATCCGTTTATGAGGTGCTGATTTGTTGCTACAGCCCTTGAATATATAT 1200
DB 1141 TTGTTATCCGTTTATGAGGTGCTGATTTGTTGCTACAGCCCTTGAATATATAT 1200
QY 1201 TCCTTTTACTATTAGAAATGCTGACTGGGGAACAGTAAAAAATTATTATAA 1254
DB 1201 TCCTTTTACTATTAGAAATGCTGACTGGGGAACAGTAAAAAATTATTATAA 1254

RESULT 2

BD087249
LOCUS
DEFINITION
1254 bp DNA linear
Hyaluronan synthase gene and utilization thereof.
PAT 27-AUG-2002

BD087249

ACCESSION

BD087249, 1 GI:26532859

VERSION

JP 2001521741-A/1.

KEYWORDS

Streptococcus dysgalactiae subsp. equisimilis

Streptococcus dysgalactiae subsp. equisimilis

Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

Streptococcus.

1 (bases 1 to 1254)

REFERENCE

1 Weigel, P.H., Kumari, K. and Deangelis, P.

Hyaluronan synthase gene and utilization thereof

Patent: JP 2001521741-A 1 13-NOV-2001;

THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA

OS Streptococcus equisimilis

PN JP 2001521741-A/1

PD 13-NOV-2001

PF 30-OCT-1998 JP 2000519083

PR 31-OCT-1997 US 60/064435; 26-OCT-1998 US 09/178851 PI

PAUL H WEIGEL, KSHAMA KUMARI, PAUL DEANGELIS

PC C12N15/09, A61K31/728, A61P43/00, C12N9/10, C12P19/04, PC

C12Q1/68//

```

PC (C12N15/09,C12R1:46),C12N15/00,C12R1:46) CC
Hyaluronan synthase gene and utilization thereof FH Key
Location/Qualifiers
FT source 1..1254
FT /organism="Streptococcus equisimilis"/
FEATURES
    source
        Location/Qualifiers
            1..1254
                /organism="Streptococcus dysgalactiae subsp. equisimilis"
                /mol_type="genomic DNA"
                /sub_species="subsp. equisimilis"
                /db_xref="taxon:119602"
BASE COUNT 335 a 240 c 261 g 418 t
ORIGIN
Query Match 100.0%; Score 1254; DB 6; Length 1254;
Best Local Similarity 100.0%; Pred. No. 1.5e-308;
Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGAACATTAAGAACCTCAATACCTGTTGGCCCTTTAGTATTTTGGGTACTGTTG 60
Db 1 ATGAGAACATTAAGAACCTCAATACCTGTTGGCCCTTTAGTATTTTGGGTACTGTTG 60
QY 61 ATTACGTCATGTTATCTCTTGGTCTAAGAGAGCTTGTCATTTATGCTTTTGG 120
Db 61 ATTACGTCATGTTATCTCTTGGTCTAAGAGAGCTTGTCATTTATGCTTTTGG 120
QY 121 CTGATAGCTTACCTATTAAGTCTCAAAATGTCCTTATCCCTTTTACAGCCATTAAGGGA 180
Db 121 CTGATAGCTTACCTATTAAGTCTCAAAATGTCCTTATCCCTTTTACAGCCATTAAGGGA 180
QY 181 AGGCTGGGCAATTAAGTCTGACCAATTAATCCCTTTATACAGAGATCGTAGTCA 240
Db 181 AGGCTGGGCAATTAAGTCTGACCAATTAATCCCTTTATACAGAGATCGTAGTCA 240
QY 241 TTGCTAGAGACCTTAAAGTGTTCAGAGCAACCTATCCCTTAGCAGAAATTTATGTT 300
Db 241 TTGCTAGAGACCTTAAAGTGTTCAGAGCAACCTATCCCTTAGCAGAAATTTATGTT 300
QY 301 GTTGACGATGAAGTGTCTGATGAGACAGGTATTAAAGCATTAAGAGTATGTCGTGAC 360
Db 301 GTTGACGATGAAGTGTCTGATGAGACAGGTATTAAAGCATTAAGAGTATGTCGTGAC 360
QY 361 ACTGTGTACCTATCAAGCAATGTCATTTTCATCGGTCTGAGAGAAATCAAGAAAGCGT 420
Db 361 ACTGTGTACCTATCAAGCAATGTCATTTTCATCGGTCTGAGAGAAATCAAGAAAGCGT 420
QY 421 CATGACAGGCGTGGGCTTTGAAGATCAGACGCTGATGCTTTTTCACGGTGTGACTCA 480
Db 421 CATGACAGGCGTGGGCTTTGAAGATCAGACGCTGATGCTTTTTCACGGTGTGACTCA 480
QY 481 GATACCTATATCTACCTGATGCTTTAGAGAGTGTGTTAAACCTTTAATGACCAACT 540
Db 481 GATACCTATATCTACCTGATGCTTTAGAGAGTGTGTTAAACCTTTAATGACCAACT 540
QY 541 GTTTTTCGTCAGCGGTGTCACCTTAATGTCAGAAATAGACAAACCAATCTCTTAACACGC 600
Db 541 GTTTTTCGTCAGCGGTGTCACCTTAATGTCAGAAATAGACAAACCAATCTCTTAACACGC 600
QY 601 TTGACAGATATTCGCTATGATATGCTTTTGGGCTTGAAGAGCTGCCCATCGGTACA 660
Db 601 TTGACAGATATTCGCTATGATATGCTTTTGGGCTTGAAGAGCTGCCCATCGGTACA 660
QY 661 GGTAATATCTTTGTCAGCTGCGCTTTAGCGTTTACAGCGGAGGTGTTGTCCT 720
Db 661 GGTAATATCTTTGTCAGCTGCGCTTTAGCGTTTACAGCGGAGGTGTTGTCCT 720
QY 721 AACATAGATAGATACCAACAGACCTTCTCGGTATTCCTGTAAGTATGGTGATGAC 780
Db 721 AACATAGATAGATACCAACAGACCTTCTCGGTATTCCTGTAAGTATGGTGATGAC 780
QY 781 AGGTGCTTGACCAACTATGCAACTGATTTAGGAAGAGCTGTTTATCAATCCACTGCTAAA 840
Db 781 AGGTGCTTGACCAACTATGCAACTGATTTAGGAAGAGCTGTTTATCAATCCACTGCTAAA 840

```

```

QY 841 TGTATTACAGATGTTCTCTGACAGATGCTACTTACTTGAAGCAGCAAAACCGCTGGAAC 900
Db 841 TGTATTACAGATGTTCTCTGACAGATGCTACTTACTTGAAGCAGCAAAACCGCTGGAAC 900
QY 901 AAGTCCTCTTTTAGAGAGTCCATTATTCTGTTAAGAAATCATGAACAATCCCTTTGTA 960
Db 901 AAGTCCTCTTTTAGAGAGTCCATTATTCTGTTAAGAAATCATGAACAATCCCTTTGTA 960
QY 961 GCCCTATGACCACTACTTGAAGTCTATGTTTATGATGCTGTTTATCTGTTGGTGGAT 1020
Db 961 GCCCTATGACCACTACTTGAAGTCTATGTTTATGATGCTGTTTATCTGTTGGTGGAT 1020
QY 1021 TTCTTTTAGCAATGTCAGAGAAATTTGATGCTGAGGCTTTCAGGCTTTTTCGCTGCTGATT 1080
Db 1021 TTCTTTTAGCAATGTCAGAGAAATTTGATGCTGAGGCTTTCAGGCTTTTTCGCTGCTGATT 1080
QY 1081 ATCTTCATTTGTTGCCCTGTCGGAACATTCATTACATGCTTAAGCACCCGCTGCTCCTC 1140
Db 1081 ATCTTCATTTGTTGCCCTGTCGGAACATTCATTACATGCTTAAGCACCCGCTGCTCCTC 1140
QY 1141 TTGTTATCTCCGTTTATGAGGCTGCTGCAATTTGTTGCTTACAGCCCTTGAATATATAT 1200
Db 1141 TTGTTATCTCCGTTTATGAGGCTGCTGCAATTTGTTGCTTACAGCCCTTGAATATATAT 1200
QY 1201 TCTCTTTTACTATTAGAAATGCTGACTGGGAACAGCTAAAAAATTATTATAA 1254
Db 1201 TCTCTTTTACTATTAGAAATGCTGACTGGGAACAGCTAAAAAATTATTATAA 1254

RESULT 3
AY173078 3552 bp DNA linear BCT 15-JAN-2003
LOCUS Streptococcus equi subsp. zooepidemicus hyaluronan acid operon,
DEFINITION Partial sequence.
ACCESSION AY173078
VERSION AY173078.1 GI:27762599
SOURCE Streptococcus equi subsp. zooepidemicus
ORGANISM Streptococcus equi subsp. zooepidemicus
REFERENCE 1 (bases 1 to 3552)
AUTHORS Wu,X., Gao,H. and Chen,J.
TITLE Molecular characterization of hyaluronan acid operon from group C
JOURNAL Streptococcus equi subsp. zooepidemicus
AUTHORS Unpublished
REFERENCE 2 (bases 1 to 3552)
AUTHORS Wu,X., Gao,H. and Chen,J.
TITLE Molecular cloning, expression, and characterization of hasC and
JOURNAL galU, two genes encoding UDP-glucose pyrophosphorylase in group C
AUTHORS Streptococcus equi subsp. zooepidemicus
REFERENCE 3 (bases 1 to 3552)
AUTHORS Wu,X. and Gao,H.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2002) Department of Biology, Tsinghua University,
Beijing 100084, China
FEATURES
    source
        Location/Qualifiers
            1..3552
                /organism="Streptococcus equi subsp. zooepidemicus"
                /mol_type="genomic DNA"
                /sub_species="zooepidemicus"
                /db_xref="taxon:40041"
            1..>3552
                /gene="hyaluronan acid operon"
            1..1254
                /gene="hasA"
            1..1254
                /gene="hasA"
            /note="required for hyaluronan acid synthesis"
            /codon_start=1
            /transl_table=11
    gene
        1..1254
            /gene="hasA"
    gene
        1..1254
            /gene="hasA"
    CDS
        1..1254
            /note="required for hyaluronan acid synthesis"
            /codon_start=1
            /transl_table=11

```

[illegible]


```
QY 301 GTTCAGATGGAAGTCTGATGACAGAGGTATTAGCGCATTTAGCACTATGTCGGTAC 360
Db 831 GTTACCATGGAAGTCTGATGACAGAGGTATTAGCGCATTTAGCACTATGTCGGTAC 890
QY 361 ACTGCTGACCTATCAAGCAATGTCACTTTTCATCGTTCAGAGAAAATCAAGAAAGCGT 420
Db 891 ACTGCTGACCTATCAAGCAATGTCACTTTTCATCGTTCAGAGAAAATCAAGAAAGCGT 950
QY 421 CATGCACAGCCCTGGCCCTTTGAAGATCAGACGCTGATGCTTTTGTGACCGTTGACTCA 480
Db 951 CATGCACAGCCCTGGCCCTTTGAAGATCAGACGCTGATGCTTTTGTGACCGTTGACTCA 1010
QY 481 GATACCTATATCTACCCCTGATGCTTTAGAGAGTGTGTAAACACCTTTAATGACCCAACT 540
Db 1011 GATACCTATATCTACCCCTGATGCTTTAGAGAGTGTGTAAACACCTTTAATGACCCAACT 1070
QY 541 GTTTTGTGCGAGCGGTACCTTAATGTCAGAAATAGACAAACCAATCTCTTAACAGCG 600
Db 1071 GTTTTGTGCGAGCGGTACCTTAATGTCAGAAATAGACAAACCAATCTCTTAACAGCG 1130
QY 601 TTGACAGATATCGCTATCAATGCTTTTGGCTTGAACGCTGCCAATCGTTACA 660
Db 1131 TTGACAGATATCGCTATCAATGCTTTTGGCTTGAACGCTGCCAATCGTTACA 1190
QY 661 GGTAAATCTCTTTTGTGCTAGGTCCTTTAGCGTTTACAGACGAGTGGTTTCTTCT 720
Db 1191 GGTAAATCTCTTTTGTGCTAGGTCCTTTAGCGTTTACAGACGAGTGGTTTCTTCT 1250
QY 721 ACATAGATAGATACATCAACAGACCTTCTGGGTATCTCTGATGATTTGGTGATGAC 780
Db 1251 ACATAGATAGATACATCAACAGACCTTCTGGGTATCTCTGATGATTTGGTGATGAC 1310
QY 781 AGGTGCTTGACCAACTATGCAACTGATTTAGGAAGACTGTTTATCAATCACTCTATAA 840
Db 1311 AGGTGCTTGACCAACTATGCAACTGATTTAGGAAGACTGTTTATCAATCACTCTATAA 1370
QY 841 TGTATTACAGATTTCTGACAGATGCTTACTTACTTGAAGCAGCAAAACCGCTGGAC 900
Db 1371 TGTATTACAGATTTCTGACAGATGCTTACTTACTTGAAGCAGCAAAACCGCTGGAC 1430
QY 901 AGTCTCTTTAGAGTCCATTTCTGTTAGAAATCATGACATCTCTTTTCTA 960
Db 1431 AGTCTCTTTAGAGTCCATTTCTGTTAGAAATCATGACATCTCTTTTCTA 1490
QY 961 GCCCATGACCACTTCTGAGTGTCTATGTTATGATGCTTGTATTTCTGCTGGTAT 1020
Db 1491 GCCCATGACCACTTCTGAGTGTCTATGTTATGATGCTTGTATTTCTGCTGGTAT 1550
QY 1021 TCTTTGTAGCAATGTCAGAAATTTGATGCTCAGGTTTGTAGCCCTTCTGCTGATT 1080
Db 1551 TCTTTGTAGCAATGTCAGAAATTTGATGCTCAGGTTTGTAGCCCTTCTGCTGATT 1610
QY 1081 ATCTTCATTTGTCCTCTGTGCGGAACATTCATTATGCTTAAGCACCCTGTCCTTC 1140
Db 1611 ATCTTCATTTGTCCTTTGTGCTGTAATTCATTATGCTTAAGCACCCTGTCCTTC 1670
QY 1141 TTGTTATCTCGTTTATGGGTGCTGCTATTTGTTGCTTACAGCCCTTGAATATAT 1200
Db 1671 TTGTTATCTCGTTTATGGGTGCTGCTATTTGTTGCTTACAGCCCTTGAATATAT 1730
QY 1201 TCTCTTTTACTATAGAAATCTGCTGGGACACGTAATAAATTTATATAA 1254
Db 1731 TCTCTTTTACTATAGAAATCTGCTGGGACACGTAATAAATTTATATAA 1784

RESULT 5
SUB242946 LOCUS SUB242946 3466 bp DNA linear BCT 06-JAN-2001
DEFINITION Streptococcus uberis hasA and hasB genes for putative hyaluronan
synthase and UDP-glucose dehydrogenase.
ACCESSION AJ242946
VERSION AJ242946.2 GI:7532762
```

hasA gene: hasB gene; hyaluronan synthase; UDP-glucose dehydrogenase.
Streptococcus uberis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.

1
Ward, P. N., Field, T. R., Ditcham, W. G., Maguin, E. and Leigh, J. A.
Identification and disruption of two discrete loci encoding hyaluronan acid capsule biosynthesis genes hasA, hasB, and hasC in Streptococcus uberis
Infect. Immun. 69 (1), 392-399 (2001)
20569205
11119529

2
Ward, P. N.
Direct Submission
Submitted (07-JUN-1999) Ward P. N., Environmental Microbiology, Institute for Animal Health, Compton Laboratory, Compton, nr Newbury, Berkshire, RG20 7NN, UNITED KINGDOM
Revised by [3]

3 (bases 1 to 3466)
Ward, P. N.
Direct Submission
Submitted (07-APR-2000) Ward P. N., Environmental Microbiology, Institute for Animal Health, Compton Laboratory, Compton, nr Newbury, Berkshire, RG20 7NN, UNITED KINGDOM
On Apr 10, 2000 this sequence version replaced gi:5441847.
Location/Qualifiers
1..3466
/organism="Streptococcus uberis"
/mol_type="genomic DNA"
/strain="01407"
/db_xref="taxon:1349"
/country="United Kingdom"
23..1276
/gene="hasA"
23..1276
/gene="hasA"
/function="putative biosynthesis of hyaluronic acid"
/codon_start=1
/transl_table=11
/product="hyaluronan synthase"
/protein_id="CAB46918.2"
/db_xref="GI:7532763"
/db_xref="SPTREMBL:Q9X9L5"
/translation="MEKLNLTFTMTFIFLWIIIGLVFVFGTKGSLTWYGIILTYLSIKMGSEFVRYPYKSGYGVKVAALIPSYNEDGVGLLETLKSVOKQTPIAEIFVIDDSVDKTKGLVEDYKLVNGFGDQVIYHQPENVGKRAQAWAFESADIVFLVDSDTIIPDALAEELATFNDPEVTAATGHLNARNRQTNLRLDIDIRDNAGVERAAQSVTGNILVCGSLTYRYSVIGNLBRYSQTFGLGVPVSGIDDRCLTNATYDLGKTYQS TARCQVDPDKFVKFIQONRMKSFRESIISVKLLATPSAVMTITEVSMFTIYVSIFSLIIGAQAQFNLIKLVAFVLVIFIVALCRNVHYMKHPFAFLSPFYGLIHLEF LQPKIYSLFTIRNATWGTKKTSK"
1301..2506
/gene="hasB"
1301..2506
/gene="hasB"
/function="putative biosynthesis of D-glucuronic acid"
/codon_start=1
/transl_table=11
/product="putative UDP-glucose dehydrogenase"
/protein_id="CAB46919.2"
/db_xref="GI:7532764"
/db_xref="SPTREMBL:Q9X9L4"
/translation="MKIIVAGSGVYGLSVLLAOKNPVTWVDIEKKVNLINOKQSP IQVDIENLYLKEKKLQLRATLDAQAERADILIIATPTNYDVEKNFPTSHVETVIE KALALNSQALLVYKSIPIGLIKMKRKQYQTRIIIFSPFERESKALKNLNLPRIIV SPEDDSEVIEAAKTFQAOLKDGSLDKVPLFMGSAEAAVLFANTYLAHRYVYF NELLDTAEKNGLRNIIIEGVDHRRIGIYHNNPFGYGGYCLPKDTKQLLAGYDGI P QSLIKAIYDSNKIRKEYIASQTLQAQLSDINVPKDATIGIYRLIMKNSDNFRRSAIK DIIDHIKSYQIINIVLYPEPMNEDFDLPIIDLSDFKAMSHIIVSNRYDLALEDVKEV YTRDIYGVN"


```

/ gene="hasA"
79..1338
/ gene="hasA"
/ function="required for hyaluronate biosynthesis"
/ note="Glc start codon"
/ codon_start=1
/ transl_except="(pos:79..81,aa:Met)
/ transl_table=11
/ product="hyaluronate synthase"
/ protein_id="RAA17725.1"
/ db_xref="GI:473747"
/ translation="MPFKKTLIVSFILISILILYLMYLFSTIVGIYVILITL
VIKLSFIYEFKPKDHYKVAAPISYNEDAESLETLKSVLAQYPLSEIYVDD
GSSNTAIOLEIEYREVIEDICRNVLNKRKHAQAWAFERSADVFLTVDSL
TYIYNALSELKSFNDEYVTAATHLNARNQTNLLTRIDRYDNAFGVERAOSL
TGNILVCSPLSIYRVELIPNLRYKQTFELGPLVSGIDRCLNYAIDLGPVYQS
TARCDVDVPEOLKSVLKOQRNWKSFRESIISVKKILSNPIVALWTFEVPWPMELI
VATGNLLFNQALQDLKLFALFSLIFIVALCRNVHVMKHPASELLSPLYGILHLFV
LQPKLISLICTIKNTGWRKKVTIFK"
1002
misc_feature
/ gene="hasA"
/ note="site of Tn916 insertion in WF62 (hyaluronate
synthase mutant)"
BASE COUNT 427 a 217 c 223 g 483 t
ORIGIN
Query Match 48.2%; Score 604.6; DB 1; Length 1350;
Best local Similarity 68.7%; Pred. No. 4.4e-143;
Matches 848; Conservative 0; Mismatches 384; Indels 3; Gaps 1;
QY 10 TTAAAAACCTCAATACGTGTGGCCCTTAGTATTTTTGGGTACTGTGTATTTACGTC 69
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
88 TTTAAAAAACTTTAATGTGTTTATCTCTTTTATCTTTTATCTATCTATCTATCTA 147
QY 70 AATGTATCTCTCTGGTCTAAAGGAGCTGTCAATTTATGCTTTTGTCTAGTAGCT 129
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
148 AATATGATCTATTGGAC---ATCACTGTAGAAATTAAGAGTAATATTAATACC 204
QY 130 TACCTATTAAGTATGCTCTATCTCTTTTACAAGCAATTTAAGGAAAGGCTGGG 189
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
205 TATCTAGTTATCAAACTTGGATTATCTTCTCTTATGAGCAATTTAAAGAAAGCCAT 264
QY 190 CAATATAAGCTGCAGCCATTTATCCCTCTATACGAGATGCTGAGTCATGCTAGAG 249
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
265 GACTATAAGCTGCTGTGTAATCTCTTATATGAAGATGCCAGTCAATATTAGAA 324
QY 250 ACCTTAAAAAGTTTCAGCAGCAAACTATCCCTTAGCAGAAATTTATGCTTGAGAT 309
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
325 ACACCTTAAAAGTTGTAGCAGACACCTATCCGTTATCAGAAATTTATTTGTTGAT 384
QY 310 GGAAGTCTGATGAGACAGGTATTAAAGCGATTAAGACTATGCGTGACACTGGTGAC 369
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
385 GGGAGTTCAACACACAGATGCCAATCAATTAATGAAGAGTATGTAATAGAGAAGTGGAT 444
QY 370 CTATCAGCAAGTCAATGTTTCATCGCTCAGAGAAAATCAAGAAAGCTCAACCCAG 429
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
445 ATTGTCGAACGTTATGTTTACCGTTCCTTGTCATAAGAAAGAAAGCCATGCTCAA 504
QY 430 GCGTGGCCCTTTGAAGATCAGACGCTGATGCTTTTTCACCGTTGACTCAGATACAT 489
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
505 GCGTGGGCAATTTGAAGATCTGACGCTGACGTTTTTTTACCGTATGATTCAGATCTAT 564
QY 490 ATCTACCCCTGATGCTTTAGAGAGTTGTTTAAACCTTTTAAACCCCACTGTTTTGCT 549
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
565 ATCTATCAAAATGCTTTAGAAGAACTCTCTAAAAGCTTCAATGATGAGACAGTTATGCT 624
QY 550 GCGAGGGTCACTTAATGTCAGAAATAGACAAACCAATCTCTTACAGCTTGACAGAT 609
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
625 GCACAGACATTTGAATGCTAGAAACAGACAAACCAATCTTATTAACGCGACTTACAGAT 684
QY 610 ATTCGCTATGATTAATGCTTTTGGGTTGACGAGTGCCTCAATCATTACAGGTAATATT 669
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
685 ATCCGTTACGATTAATGCTTTGGGGTGGAGCGTGTCTCAATCATTACAGGTAATATT 744

```

```

QY 670 CTGTGTTGCTAGTCGCTTAGCGTTTACAGACGCGAGGTGGTGTTCCTTAACATAGAT 729
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
745 TTAGTTTGTCTAGGACCAATTTAGTATTTTACGAGTGAAGTGAATTTTCTTAACTTAGAG 804
QY 730 AGATACATCAACAGACGCTTCTGGGTATTCCTCTAGTATTGGTGACGAGCTGCTTG 789
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
805 CGTATTAATAATCAACATTTCTTAGTATTACCTCTAGCATTTGGGATGATCATGTTTA 864
QY 790 ACCAATATGCAACTGATTTAGGAAGACTGTTTATCAATCCACTGCTAAATGATTAACA 849
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
865 ACAATATATGCTATTGATTTAGGACGCACTGCTTACCAATCAACAGTAGATGATAC 924
QY 850 GATGTTCTGCAAGATGCTACTTACTTGAAGCAGCAAAACCGCTGGAACAAGTCCCTC 909
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
925 GATGTACCTTTCCCAATTAAGAGTTATTTAAGCAACAAATCGATGGAATAAATCTTT 984
QY 910 TTTAGAGAGTCCATATTTCTTGAAGAAATCATGAACAATCCCTTTGTAGCCCTATGG 969
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
985 TTTAGAGAACTATTTCTGTTAAAAAATCTTTCTAATCCCATCGTTGCTTTATGG 1044
QY 970 ACCATATGAGTGTCTATGTTTATGATGCTGTTTATTTCTGTGTGGTGAATTTCTTTGA 1029
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1045 ACTATTTGGAAGTCTGTTATGTTGATGTTGATGTCGAATTTGGCAATTTGGAAATCTTT 1104
QY 1030 GGCAATGTCAGAGAAATTTGATTTGGCTCAGGCTTTTAGCCCTTTCTGTCATATCTTCAT 1089
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1105 AATCAAGCTATTTCAATTTAGACCTTATTAACCTTTTGGCTTTTATCAATCATCTTTATC 1164
QY 1090 GTTGCCCTGTGTGCGGAACATTCATTAATGCTTAAGCACCCTGCTGCTCTTCTTGTATCT 1149
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1165 GTTGCTTTATGCTGAATGTTTATGTTATGTTGTCGAATCTGCTGCTGCTGCTGCTGCTGCT 1224
QY 1150 CCGTTTATGAGGCTGCTGCAATTTGTTGCTCTACAGCCCTTGAATTAATTTCTCTTTT 1209
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1225 CCTCTGATGAATTTACACTGTTTGTCTTACAGCCCTTAAACTTTATCTTTTATGCT 1284
QY 1210 ACTATTAGAATGCTGACTGGGGAACAGTAAAAA 1244
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1285 ACCATTAAAAATACGAATGGGGAACAGTAATAA 1319
STRHA5 1440 bp DNA linear BCT 09-MAY-1994
DEFINITION Streptococcus pyogenes hyaluronan synthase (hasA) gene, complete cds
ACCESSION L20853
VERSION L20853.1 GI:410042
KEYWORDS hasA gene; hyaluronan synthase; hyaluronate; hyaluronic acid.
SOURCE Streptococcus pyogenes
ORGANISM Streptococcus pyogenes
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 1440)
AUTHORS DeAngelis,P.L., Papaconstantinou,J. and Weigel,P.H.
TITLE Molecular cloning, identification, and sequence of the hyaluronan synthase gene from group A Streptococcus pyogenes
J. Biol. Chem. 268 (26), 19181-19184 (1993)
MEDLINE 93374890
PUBMED 8366070
COMMENT Original
FEATURES
Source
text: Streptococcus pyogenes (strain 543) DNA.
Location/Qualifiers
1..1440
/organism="Streptococcus pyogenes"
/strain="543"
/db_xref="taxon:1314"
109..1380
/gene="hasA"
109..114
/gene="hasA"
/ note="possible ribosome binding site at -109 (upstream of atg start codon) near alternate start codon (GTG) at

```


Db 1147 ATCAAGTATCAATAGACCTTATTAACCTTTTGGCTTTTATCCATCATCTTATC 1206

QY 1090 GTTGGCCCTGTCGGAACATTCATTACATGCTTAAGCACCAGCTGCTCTCTGTTATCT 1149

Db 1207 GTTGGCTTATGCTGATGCTTATTAATGCTCAACATCTGCTAGTTTGTATCT 1266

QY 1150 CCGTTTATGGGGTCTGCAATTTGTTGCTCAGCCCTGGAATATATATCTCTTTT 1209

Db 1267 CCGTGTATGAATATATACACTGTTGCTTACAGCCCTTAAACCTTTATCTTATGC 1326

QY 1210 ACTATTAGAAATGCTGACTGGGGAACAGTAAAA 1244

Db 1327 ACATTAATAATAGGAATGGGGAACAGTAAAA 1361

RESULT 8

AR232201

LOCUS AR232201 1512 bp DNA linear PAT 20-DEC-2002

DEFINITION Sequence 1 from patent US 6455304.

ACCESSION AR232201

VERSION AR232201.1 GI:27274189

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1512)

AUTHORS Weigel,P.H., DeAngelis,P.L. and Papaconstantinou,J.

TITLE Hyaluronate synthase gene and uses thereof

JOURNAL Patent: US 6455304-A 1 24-SEP-2002;

FEATURES

Location/Qualifiers

1..1512

source /organism="unknown"

BASE COUNT 484 a 227 c 245 g 556 t

ORIGIN

Query Match 48.2%; Score 604.6; DB 6; Length 1512;

Best Local Similarity 68.7%; Pred. No. 4.4e-143;

Matches 848; Conservative 0; Mismatches 384; Indels 3; Gaps 1;

QY 10 TTAATAAACCTCAATACCTGTTGGCTTTAGTATTTTGGGTAATGTTTACGTC 69

Db 202 TTTAAAAAATAATTAATGTTTATCTTTTATCTTTTATCTTTTATCTTTTATCTA 261

QY 70 AATGTTTATCTTTTGGCTTAAGGAGCTTCTCAATTTATGCTTTTGTGATAGCT 129

Db 262 AATATGATCTTTTGGAACTTTTATCTTTTATCTTTTATCTTTTATCTTTTATCTA 318

QY 130 TACCTATTAAGTAAATGCTTTTATCTTTTATCTTTTATCTTTTATCTTTTATCT 189

Db 319 TATCTATGTTTATCAACTTGGATATCTTTTATCTTTTATCTTTTATCTTTTATCTA 378

QY 190 CAATATAGTTTGCACCAATTTATCTTTTATCTTTTATCTTTTATCTTTTATCTA 249

Db 379 GACTATAAGTTGCTGCTTAATCTTTTATCTTTTATCTTTTATCTTTTATCTTTA 438

QY 250 ACCTTAAAGTGTTCAGCAGCAAACTATCCCTTATCTTTTATCTTTTATCTTTTATCT 309

Db 439 ACACCTTAAAGTGTTCAGCAGCAAACTATCCCTTATCTTTTATCTTTTATCTTTTATCT 498

QY 310 GGAAGTCTGATGAGACAGGATTAAGGCAATTAAGGCAATTAAGGCAATTAAGGCA 369

Db 499 GGGAGTTCAACACAGATGCAATTAAGGCAATTAAGGCAATTAAGGCAATTAAGGCA 558

QY 370 CTATCAACATGCTATCTTCACTGCTGAGCAAAATCAAGCAAAATCAAGCAAAATCA 429

Db 559 ATTGTGCAAACTGATCTTCACTGCTGAGCAAAATCAAGCAAAATCAAGCAAAATCA 618

QY 430 GCGTGGCTTTGAAAGATCAGACGCTGATGCTTTTATCTTTTATCTTTTATCTTTTATCT 489

Db 619 GCGTGGCAATTTGAAAGATCAGACGCTGATGCTTTTATCTTTTATCTTTTATCTTTATCT 678

QY 490 ATCTACCTCTGATGCTTTAGAGGATGTTTAAAAACCTTTTATGACCAACTGTTTGTCT 549

Db 679 ATCTATCAAAATGCTTTAGAGAACTCTTAAAGCTTCAATGATGAGACAGTTTATGCT 738

QY 550 GCGAGGGTCACTTAAATGCTAGAAATAGACAAACAACTCTCTTTAACAGGTTGACAGAT 609

Db 739 GCAACAGGACATTTGAATGCTAGAAACAGACAACTATCTTAAACGCACTTACAGAT 798

QY 610 ATTGCGTATGATAATGCTTTTGGCTTTGAACGAGTGCCTCAATCCGTTACAGGTAATATC 669

Db 799 ATCCGTTAGTAATGCTTTTGGGGTGGAGCGTCTCTCAATCAATTAACAGGTAATATC 858

QY 670 CTGTTTGTCTCAGTCCGCTTACAGTTTACAGACGAGGTTGTTTCTTAAACATAGAT 729

Db 859 TTAGTTTGTCTCAGGACCACTTGAATTTATCGAGTGAAGTATTTCTTAACTTAGAG 918

QY 730 AGATACATCAACAGACCTTCCCTGGGTATCTCTGTAAGTATTTGGTGATGACAGGTCGTTG 789

Db 919 CGCTATTAATAATCAAACTTCTAGGTTTACCTGTTAGCATTGGGGATGATCGATGTTA 978

QY 790 ACCAATATGCAACTGATTTAGGAAAGACTGTTTATCAATCCACTGCTAAATGTTATTACA 849

Db 979 ACAATATATGCTATTGATTTAGGACCGCTGCTTACCAATCAGACGATGATGATGATCT 1038

QY 850 GATGTTCTCTGACAAGTGTCTACTTACTTGAAGCAGCAAAACCGCTGGACAACTCTTC 909

Db 1039 GATGTTCTCTTCCAAATTAAGAGTATTTAAAGCAACAAATCGATGCAATTAATCTTTT 1098

QY 910 TTTAGAGAGTCCATTTATCTGTTTAAAGAAATCATGACAACTCTTTTGTAGCCCTATGG 969

Db 1099 TTTAGAGAACTATTTATCTGTTTAAAGAAATCTTTCTTAACTCCATGTTGGCTATGG 1158

QY 970 ACCATACCTTGGAGTGTCTATGTTTATGATGCTTTTATCTTGTGCTGATTTCTTTGTA 1029

Db 1159 ACTATTTTCCGAAGTCTTATGTTTATGATGTTGATTTCTGCAATTTGGAACTTTTGT 1218

QY 1030 GGCAATGCTCAGAGATTTGATTTGGCTCAGGTTTAACTCTTCTGCTGATTAATCTTCAAT 1089

Db 1219 AATCAAGCTATTTCAATTAGACCTTATTAACCTTTTTCCTTTTATCCATCATCTTTATC 1278

QY 1090 GTTGGCCCTGTCGGAACATTCATTACATGCTTAAGCACCCTGCTCTCTTCTTCTTATCT 1149

Db 1279 GTTGGCTTATGCTGTAATGTTTATTTATGCTCAACATCTCTGCTAGTTTGTGTTATCT 1338

QY 1150 CCGTTTATGGGGTCTGCTGCAATTTGTTGCTTACAGCCCTTGAATATATATCTTCTTTTT 1209

Db 1339 CCTCTATGGAATATATACACTTTGTTGCTTACAGCCCTTAAACTTTATCTTCTTATGC 1398

QY 1210 ACTATTAGAAATGCTGCTGAGGCAACAGTAAAA 1244

Db 1399 ACCATTAATAATGCGAATGGGGAACAGTAAAA 1433

RESULT 9

AE010123

LOCUS AE010123 10851 bp DNA linear BCF 03-APR-2002

DEFINITION Streptococcus pyogenes strain MGAS8232, section 171 of 173 of the complete genome.

ACCESSION AE010123 AE009949

VERSION AE010123.1 GI:19749282

KEYWORDS

SOURCE Streptococcus pyogenes MGAS8232

ORGANISM Streptococcus pyogenes MGAS8232

Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.

REFERENCE 1 (bases 1 to 10851)

AUTHORS Smoot,J.C., Barbian,K.D., Van Gompel,J.J., Smoot,L.M., Chaussee,M.S., Sylva,G.L., Sturdevant,D.E., Ricklefs,S.M., Porcella,S.F., Parkins,L.D., Beres,S.B., Campbell,D.S., Smith,T.M., Zhang,Q., Kapur,V., Daly,J.A., Veasy,L.G. and Musser,J.M.

TITLE Genome sequence and comparative microarray analysis of serotype M18 group A Streptococcus strains associated with acute rheumatic fever outbreaks

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (7), 4668-4673 (2002)

MEDLINE 21927593

11917108

PUBMED
REFERENCE
AUTHORS

2 (bases 1 to 10851)
Smoot,J.C., Barbican,K.D., Van Gompel,J.J., Smoot,L.M.,
Chaussee,M.S., Sylva,G.I., Sturdevant,D.E., Rickierts,S.M.,
Forcella,S.F., Parkins,L.D., Beres,S.B., Campbell,D.S., Smith,T.M.,
Zhang,Q., Kapur,V., Daly,J.A., Veasy,L.George. and Musser,J.M.
Direct Submission

TITLE
JOURNAL
Submitted (31-JAN-2002) Laboratory of Human Bacterial
Pathogenesis/Rocky Mountain Laboratories/NIH, 903 S. 4th St.,
Hamilton, MT 59840, USA

FEATURES
source

Location/Qualifiers
1. 10851

/organism="Streptococcus pyogenes MGAS8232"
/mol_type="genomic DNA"
/strain="MGAS8232"
/db_xref="taxon:186103"
complement(238..852)
/gene="spyM18_2226"
complement(238..852)
/gene="spyM18_2226"
/notes="best blastp match gb|AAK34820.1| (AE006636)
hypothetical protein [Streptococcus pyogenes M1 GAS]"
/codon_start=1

/transl_table=11
/product="hypothetical protein"

/protein_id="AAL98659.1"
/db_xref="GI:19749283"

/translation="MFKENLKORFENGLVALALITLITAFESKNADPKSAKKS
ESKWTIDAPKNHATIKESKEKASIASEPITVENSVAPTVERAPVQOEVTQ
TVQCVSSVAINNNVLSNGTAGIVGQAQAQAATVGPSTWEHIIARENSGNPN
AANASGSLFQTPMGWGTATVEDQVNAALKAYSAGLSAMGY"

complement(1223..2023)
/gene="spyM18_2228"
complement(1223..2023)

/gene="spyM18_2228"
/notes="best blastp match gb|AAK34821.1| (AE006636)
conserved hypothetical protein [Streptococcus pyogenes M1
GAS]"

/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"

/protein_id="AAL98660.1"
/db_xref="GI:19749284"

/translation="MDKLILGRYPGDSLIHRLDPRSKLLMIYIVIFWANNVNT
LLMTTLAVFLSKILSFLGVKPMIGILFTILFQFSGQGVIFSWFEISIT
DLGISAILIFNRVLIIFSTLLITLTPLSDAVESLLKPLTRFKVPAHETGLM
SLSRVPTLMDTTRIMNAQRAGVDREGNLQKVKSIIPIILPLFASFKRADAL
AIAMEARGVQSGEGRTKYRQLDQKSLAIGVISLGLLILFLKTSI"

complement(2016..2858)
/gene="spyM18_2229"
complement(2016..2858)

/gene="spyM18_2229"
/notes="best blastp match gb|AAK34822.1| (AE006636)
putative ABC transporter (ATP-binding protein)

[Streptococcus pyogenes M1 GAS]"
/codon_start=1

/transl_table=11
/product="putative ABC transporter (ATP-binding protein)"

/protein_id="AAL98661.1"
/db_xref="GI:19749285"

/translation="MSINLQNSVTVQAGTPPEGRALENLNDILDGSYTAFIGHGS
GKSTIQMLNGLHVPTGTIVSDQDITNHSKNEIKSIRKHVGLVGFQFPESQLEET
VLKDVAFQNGVSPFEALAREKALVAGISNLFERNFELSGQMRRAVAGIL
AMQPKVLVDEPTAGLDPKGRKELMTIFKIHQSGMTIVLVTHLMDADVADFVYL
DKQKILSGKPTTIQQVSLLEKKQLGVKVTKLAQRVDRGPIPSLPILEELREV
LKHC"

complement(2834..3724)
/gene="spyM18_2230"
complement(2834..3724)

/gene="spyM18_2230"
/notes="best blastp match gb|AAC61484.1| (AF082738) ABC
transporter (ATP-binding protein) [Streptococcus
pyogenes]"

/codon_start=1
/transl_table=11
/product="hypothetical protein"

/codon_start=1
/transl_table=11
/product="ABC transporter (ATP-binding protein)"
/protein_id="AAL98662.1"
/db_xref="GI:19749286"
/translation="MIILKEAFLRLNNMSAIELKVTYFNKHQOEKPTLDVGS
FHVQGGWLSIIGHNGSGKSTIRLDIGLEPESGIIVDGLLTITNWEIRHKIGM
VQNPDDQFVGAVVEDDVAFGLENGKIAHEDIKERVNHALEVMQNFKEKPARLIG
GQKRVAGAVAKFKIILIDEATSMIDPKRLEIKTIKNIPDDIQTVISITHDL
DEVALSURVLMDKQGVSTPQELFARGDELLOGLDIPFTTSVVVQMLQEGEYPTD
YGYLKELENQLCLISKM"

gene

CDS

complement(3675..4217)
/gene="spyM18_2231"
complement(3675..4217)

/gene="spyM18_2231"
/notes="best blastp match gb|AAK34824.1| (AE006636)
phosphatidylglycerophosphate synthase [Streptococcus
pyogenes M1 GAS]"
/codon_start=1

/transl_table=11
/product="phosphatidylglycerophosphate synthase"

/protein_id="AAL98663.1"
/db_xref="GI:19749287"

/translation="MIKKENIPNLLTLVRIAMIPFLEFITSNNKYGVHIFAIVFAI
ASFTDYLGDYLRKRWVSNFKGADPLADKMLVMSAFIMLVGLVPAWVSAVLICR
ELAVTGRLRLVFTGGKVLAAAMLGKIKTKATQMLSTILLCHWIFLGNLLCIAFFT
IYSGVDYFKGASFLKNDKEK"

complement(4231..5256)
/gene="spyM18_2232"
complement(4231..5256)

/gene="spyM18_2232"
/notes="best blastp match gb|AAC61482.1| (AF082738) unknown
[Streptococcus pyogenes]"

/codon_start=1
/transl_table=11

/product="hypothetical protein"

/protein_id="AAL98664.1"
/db_xref="GI:19749288"

/translation="MSEKSLGVLRESRVGNKNTLDDIELATGISSHYLLAMELDOEK
LIPEKFDQELKEYADIVELDFNLKRRYQVNSKSDISKSVTEIVEKLSKRQIQ
EKLSHQSGDGGQPLEVSKNNTLVIPKTVGLDLSKPKTISPVQSKDNRLLPKK
NYSRMRNSTDSKTSFLPAILITAFATISVILGVQWQFQSKRAEVALPK
ASEQSTGLKNNDNRKTQVTEGSGNYLIATVKETVDIVSVLSEQQSSWISLNN
SEIGGGGTTITQSPYATATPAETESLTLTGVTNGVSVNTIDGQPVLDLSPLTSDLS
YITFKIO"

complement(5306..6595)
/gene="spyM18_2233"
complement(5306..6595)

/gene="spyM18_2233"
/notes="best blastp match gb|AAC61481.1| (AF082738) unknown
[Streptococcus pyogenes]"

/codon_start=1
/transl_table=11

/product="hypothetical protein"

/protein_id="AAL98665.1"
/db_xref="GI:19749289"

/translation="MTKLKYNPNIDEVLYVKLENGLTYYFKIKIGFLEKTAFLTV
GFGSDKLIVDSDRPAQTAHFLEKLFDESGDILSKTQLGAEATNAFTTNG
TSFEFSASFQONLELQYVLSANITDESYSREKKLIGOEIDMYDDADYRAYSGI
LQNLKPTLSANIDAGRESQTKTILLETHTHTFYQPTNLSLFVIGDIDDEFLLA
IQRTPTLSPDRKRVTVDPHLYPIKSSSDVMDVTTAKLVGFRGIVLTLTLSLT
YRIALKUFLKLGWTSKYHTLYEDGIDDSDFDVEIHNFQFVLIISLTPPEIAM
SNYIKQITKLSKEFTNEHLNLLKEMYGDFIQSLDSIEHLTHOFSLYSDSKET
YFDIPKTIERTLKVDTTIGKATFEKADASDFTVFPK"

complement(6597..7841)
/gene="spyM18_2234"
complement(6597..7841)

/gene="spyM18_2234"
/notes="best blastp match gb|AAC61480.1| (AF082738) unknown
[Streptococcus pyogenes]"

/codon_start=1
/transl_table=11

/product="hypothetical protein"

Db	8714	ATCTATCCAAATGCCCTTAGAAGAACTCCTAAAAAGCTTCAATGATGACAGAGTTTATGCT	8777
QY	550	GGCAGGGTGACCTTAAATGTCAGAAATAGCAAAACCAATGCTCTTAACACGCTTGACAGAT	609
Db	8774	GCACAGACATTTGAATGCTTAGAACACAGAAACTATCTATTACCGGACTTACAGAT	8833
QY	610	ATTCGCTATGATATGCTTTTGGGTTGAACAGAGTCCCAATCCGTTTACAGGTAATATC	669
Db	8834	ATCCGTTACGATATCCCTTTGGGTGGAGCGTGTCTCAATCATTAACAGGTAATAT	8893
QY	670	CTTGTTCCTCAGGTCGCTTAGGTTTACAGACGGAGGTGGTTCTTCCTACATAGAT	729
Db	8894	TTAGTTTCCTCAGACCAATGATGATTTTATCAGCTGGAGTGATATTCTCTACTTACAG	8953
QY	730	AGATACATCAACAGACCTTCCTGGGTATTCCTGTAAGTATTGGTGATGACAGGTGCTTG	789
Db	8954	CGCTATAAAATCAAAATCTCTAGGTTTACCTGTAGCAATGGGATGATCGATGTTA	9013
QY	790	ACCAACTATGCACATGATTAGGAAGACGTGTTATCAATCCACTGCTTAATGATATACA	8494
Db	9014	ACAAATATGCTATTGATTTAGGACGCACTGCTACCAATCAACAGCTAGATGTGATAT	9073
QY	850	GATGTTCTCAGCAAGATGCTACTTACTTGAAGCAGCAAAACCGCTGGAAACGTCCTTC	909
Db	9074	GATGTACCTTTCCAAATTAATAAGTTATTTAAAGCAACAAAATCGATGAATAATCTTT	9133
QY	910	TTTACAGAGCCATATTTCTGTTAGAAAATCATGAACATCCCTTTTGTAGCCCTATGG	969
Db	9134	TTTACAGAACTATATTTCTGTTTAAAAAAATCTTCTAATCCCACTGGTGCCTATATGG	9193
QY	970	ACCACTACTGAGGTGCTATGCTTTATGATGCTGTTGTTATTTCTGTGGTGGATTTCTTTGTA	1029
Db	9194	ACTATTTCCGAAGTCGTTATGTTTATGATGTTGATCTCGCAATTTGGGAATCTTTGTT	9253
QY	1030	GGCAATGTCAGAAATTTGATGGCTCAGGGTTTACGTTTCTGGTGATATATCTCAAT	1089
Db	9254	ATCAAGCTATTCAAATTAGACCTTATTAAACATTTTCCCTTTTATCCATCATCTTTATC	9313
QY	1090	GTTGCCCTGTGTCGGAACATTCATACATGCTTAAAGCACCGCTGCTCCTCTTGGTTATCT	1149
Db	9314	GTTGCTTTATGTCGTAATGTTTCATATATGGTCAACATCTGCTGCTTTTTCGTTATCT	9373
QY	1150	CGTTTATGGGTGCTGCATTTGTTGCTCAGAGCCCTGAAATATATATCTCTTTTT	1209
Db	9374	CCTCTGTATGGAATATACACTTGTGTGCTTACAGCCCTAAAACTTTATCTTATGTC	9433
QY	1210	ACTATTAGAATCTGACTGGGGAACACGTAAAA	1244
Db	9434	ACATTAATAATACGAATGGGAACACGTAAAA	9468
RESULT 10			
AE006637			
LOCUS	AE006637	9844 bp DNA linear	BCT 01-JUN-2001
DEFINITION	Streptococcus pyogenes M1 GAS strain SF370, section 166 of 167 of the complete genome.		
ACCESSION	AE006637		
VERSION	AE004092		
KEYWORDS	AE006637.1 GI:13623172		
SOURCE			
ORGANISM	Streptococcus pyogenes M1 GAS		
	Streptococcus pyogenes M1 GAS		
	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.		
REFERENCE	1 (bases 1 to 9844)		
AUTHORS	Perrettu, J.J., McShan, W.M., Adic, D., Savic, D., Savic, G., Lyon, K., Ferretti, J.C., Sezaev, S., Surorov, A.N., Kenton, S., Lai, H., Lin, S., Qian, Y., Jia, H.G., Najjar, F.Z., Ren, Q., Zhu, H., Song, L., White, J., Yuan, X., Clifton, S.W., Roe, B.A. and McLaughlin, R.E.		
TITLE	Complete genome sequence of an M1 strain of Streptococcus pyogenes		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 98 (8), 4658-4663 (2001)		
MEDLINE	21132684		
PUBMED	11296296		
REFERENCE	2 (bases 1 to 9844)		

AUTHORS

Ferretti, J.J., McShan, W.M., Adijic, D., Savic, D., Savic, G., Lyon, K., Primeaux, C., Sezate, S.S., Surorov, A.N., Kenton, S., Lai, H., Lin, S., Qian, Y., Jia, H.G., Najjar, F.Z., Ren, Q., Zhu, H., Song, L., White, J., Yuan, X., Clifton, S.W., Roe, B.A. and McLaughlin, R.E.

TITLE

Direct Submission

JOURNAL

Submitted (10-Apr-2001) Department of Microbiology and Immunology, University of Oklahoma Health Sciences Center, 940 SL Young Blvd, Oklahoma City, OK 73104, USA

FEATURES

Location/Qualifiers

1..3944

/organism="Streptococcus pyogenes M1 GAS"

/mol_type="genomic DNA"

/strain="SF370"

/serotype="M1"

/db_xref="taxon:160490"

246..1505

/gene="hasA"

/note="synonym: SPY2200"

246..1505

/gene="hasA"

/note="Best Blastp hit = pir|A53100 hyaluronate synthase A (HasA) - Streptococcus sp. (group A)"

>gi|473747|gb|AA17725.1| (L21187) hyaluronate synthase

[Streptococcus pyogenes]"

/codon_start=1

/transl_table=11

/product="hyaluronate synthase"

/protein_id="AAK34828.1"

/db_xref="GI:13623173"

/translation="MPIFKATLIVLSPFLISILYINMYLGSTGVIGVILIVYL
VILGLISFLYEPFKPHDYKAAVIPSYNEDAESLETLKSLVLAQYPISEIYVDD
GSNTADQLIEYVREVDICNVIVHRSVLRKGRHAQAWAFSDADYVTDSD
TYVPALELLKSFNDYVIAATGLNARNQTNLITLDIRVDNAGVRAAQS
TGNILSGPLISYREVLIPLNERYKNOFLGLPVSIGDRLCYAIDIGRTVQS
TARCDDPFLKSLYLAQNRWKSFEKESIIISVKILSNPIVALTWIEVWPMMLI
VAQNLLENQALQDLKURFALISLIFVLCRCNVHINKHFASELLSPLGLIHLFV
LQPKLISQTLKNTKMTGTRKKVTIFK"

1541..2749

/gene="hasB"

/note="synonym: SPY2201"

1541..2749

/gene="hasB"

/EC_number="1.1.1.22"

/note="Best Blastp hit = sp|Q07172|UDG_STRPY UDP-GLUCOSE
6-DEHYDROGENASE (UDP-GLC DEHYDROGENASE) (UDP-GLCDH)

(UDPGDH) >gi|1075776|pir|A46089 UDPglucose

>gi|8569328|pdb|1DL1|A Chain A, The First Structure Of

UDP-Glucose Dehydrogenase (Udpgh) Reveals The Catalytic

Residues Necessary For The Two-Fold Oxidation

>gi|295189|gb|AA26899.1| (L08444) UDP-glucose hydrogenase

[Streptococcus pyogenes]"

/codon_start=1

/transl_table=11

/product="UDP-glucose 6-dehydrogenase"

/protein_id="AAK34829.1"

/db_xref="GI:13623174"

/translation="MKIAGSGYVGLSGVLSLQNEVTIVDLPKSVKINNGSLP
IQEYEVYKLSKADLSKAKEAEVLIATPYNRNRYEDTOHETVVK
EVLVNSHATLILKSTIPICFITEMKQFODRIIESPEFLSKALVDNLYPSRIIV
SCENDSPKVRADAEKFAILLKSAKKNNPVILMGASAEAVKLFANTYALAKYVIF
QTLTAEKRLNLSHMIIGISYDRIYGRHMYNPNFSGYCYLPKDTKQLLYANNIP
QTLEIVSNRNKRSYIAQIINLEERSPVKVGVYRLMKNSDNFESAIDKV
IDILKSKDIKIIIEPMLNKLESDQSVLVNDLENFKKQANIIVTNRYDNLQDVNKK
VYSROIFNRD"

2931..3845

/gene="hasC"

/note="synonym: SPY2202"

2931..3845

/gene="hasC"

/EC_number="2.7.7.9"

/note="Best Blastp hit = sp|O54713|HASC_STRPY
UTP-GLUCOSE-1-PHOSPHATE URIDYLTRANSFERASE (UDP-GLUCOSE

PYROPHOSPHORYLASE) (UDPGP) (ALPHA-D-GLUCOSYL-1-PHOSPHATE
URIDYLTRANSFERASE) (URIDINE DIPHOSPHOGLUCOSE
PYROPHOSPHORYLASE) >gi|995767|gb|AAA91810.1| (U03452)
UDP-glucose pyrophosphorylase [Streptococcus pyogenes]"

/codon_start=1

/transl_table=11

/product="UDP-glucose pyrophosphorylase"

/protein_id="AAK34830.1"

/db_xref="GI:13623175"

/translation="MTKVKKAIIPAGLGTRELPAKAKEMLPIDVPTQFIIVEE
ALKSGIEILVITGKAKRSIEDHFSNFEYLNQAKNELKLKLVDETTAINLHFTIR
QSHRGLGDAVLAQAKAFVGNPPVFMVGLMDLITNASKAPLTKALMEYDKTHASTIR
AMKVPHEDEVSSYVIAQGVKAVGLYSDVTEVKEPOEDAPSDLAIGRYLLPETF
GILEROTPGAGNEVOLTDAITLANKTORVFAREFKGNRYDVGKFGFKMTSIDVALSH
POVKEDLNKYLKGLKALEKSKVPYTHSK"

4153..4566

/gene="SPY2203"

4153..4566

/gene="SPY2203"

/note="Best Blastp hit = pir|JC4754 hypothetical 13.6k

protein - Lactococcus lactis >gi|1107709|emb|CAA61548.1|

(X89367) orf121 [Lactococcus lactis]"

/codon_start=1

/transl_table=11

/product="conserved hypothetical protein"

/protein_id="AAK34831.1"

/db_xref="GI:13623176"

/translation="MYIKLFTFTTQALLKELGIQSGAIGKFLAETTVLFNGEBS
RRKKIRVGDKLSIPDDQDLITIVPEEQEEQFAEEMAEKTRVAALYKQMNQANKK
TSSKNNRQSTTKSLRATKTKTPAPVFPFGI"

4568..5674

/gene="recF"

/note="synonym: SPY2204"

4568..5674

/gene="recF"

/note="Best Blastp hit = sp|P49999|REFC_STRPY RECF PROTEIN

>gi|1075773|pir|JC4077 recf protein - Streptococcus

pyogenes >gi|533080|gb|AAA85783.1| (U07342) Recf protein

[Streptococcus pyogenes]"

/codon_start=1

/transl_table=11

/product="Recf protein"

/protein_id="AAK34832.1"

/db_xref="GI:13623177"

/translation="MKIKELKHRYNDHLLASFSSGNVFNIGNAOKTNFLAIV
FLSLTRSHRRADKELIHFDHSTVSLTGKIORISGTVLEINLSDKGRVTKINALKA
KLSDYCTMNVLFAPEDQLYKCAPSLRKRKIDIDLQIKPVLSELSHNVHLKOR
NSYIAKSAQQIDRAFLAVLDEQLASIGARVNEHIDF INALEKANTHHQAISNGLESU
SLUSYGVSVFDDKNTIYQOFLQLEKHNQKDFRANTSGPHRDELAFYINGMANFA
SQGHRSLIILSKMAEVSILMKALACDNPILLDDVYSELNDTRQTKLETIKENVQT
FTTTSLDLSQLEPIGFHFHTKGTQVQDSDIH"

complement(5729..6562)

/gene="SPY2205"

complement(5729..6562)

/gene="SPY2205"

/note="Best Blastp hit = sp|P40419|YCXE_BACME HYPOTHETICAL

30.5 KD PROTEIN IN GDHI 5'REGION (ORF 2)

>gi|801221|pir|J03084 hypothetical 30.5K chain - Bacillus

megaterium >gi|216267|dbj|BAAL4098.1| (D90043) ORF2

[Bacillus megaterium] >gi|228656|pirf|1808263B gdh1

upstream ORF2 [Bacillus megaterium]"

/codon_start=1

/transl_table=11

/product="conserved hypothetical protein"

/protein_id="AAK34833.1"

/db_xref="GI:13623178"

/translation="MFTWGSIGFVSNNKIGKPSQOTLGMTFGALLFSLAVLIVRPEM
TLQLWFLIGLGFINSIGTQMGVSVANPLSSQSLVLSGLIGLVFVHEWT
RPMOFVGSLLALLIVGYFSSKODDANAQVNLHNEPSKFRALTYSTIGVYAVL
FNIMKFEVLSTLPMAYGVLAGITPMSEFKISIDQVVKNSVVGILMGIGNIFMLLA
ASKAGLAIAFSQSGLAISIVGGLFGLGETKTKEMRWVTVGIICFVIGAILGVVK
S"

complement(6794..8275)

gene

[illegible]

AUTHORS

Beres,S.B., Sylva,G.L., Barbian,K.D., Lei,B., Hoff,J.S.,
Mammarella,N.D., Liu,M.-Y., Smoot,J.C., Porcella,S.F.,
Parkins,L.D., McCormick,J.K., Leung,D.Y.M., Schlievert,P.M. and
Musser,J.M. Mission

JOURNAL

Submitted (14-JUN-2002) Laboratory of Human-Bacterial Pathogenesis,
Rocky Mountain Laboratories, NIAID, NIH, 903 South Fourth St.,
Hamilton, MT 59840, USA

FEATURES

Location/Qualifiers

1..53896

/organism="Streptococcus pyogenes MGAS315"

/mol_type="genomic DNA"

/strain="MGAS315"

/serotype="M3"

/db_xref="taxon:198466"

/note="group: A"

complement(200..1480)

/gene="hiss"

/note="synonym: SpyM3_1815"

complement(200..1480)

/gene="hiss"

/note="best blastp hit: gb|AAL98633.1| (AE010120) putative
histidine-tRNA ligase [Streptococcus pyogenes MGAS8232]"

/codon_start=1

/transl_table=11

/product="putative histidyl-tRNA synthetase"

/protein_id="AAM80422.1"

/db_xref="GI:21905567"

/translation="MKLQKPGTODILPCDAKQWYVESVARDTSQNYGEITPMF
EHEVLSRVGTDITDVKEMDFDKGDRHTLRPEGTAPVRSVFNKULFAPEVOK
PVKLYIGSMFYERPGQRLREFHQVEGFGAANPATDVTETAMAYHLEFKIGIND
VTLHLSIGSPERSAVQALIDITPMRDQSKDSQRDLNPLRVLDSKEKEDIA
VERAPSLIDYDEESQAHFAVKALEALDIPYIDTNMVRGLDYNHTIFEITFIVE
GSOLITCAGRDYSLVGFPGTFCGFLGRLLRWIEKQGITLPIETEMDVILA
VLGDGANSKALEVQAIRROGTAERDYLGRKIRAKFSADTFKAKLVWLTGESEVEA
GKAVIKNRSRQVEVSEFEDMTNPNANISEQLLS"

gene

1700..1882

/gene="rpmf"

/note="synonym: SpyM3_1816"

1700..1882

/gene="rpmf"

/note="best blastp hit: gb|AAL98635.1| (AE010120) 50S
ribosomal protein L32 [Streptococcus pyogenes MGAS8232]"

/codon_start=1

/transl_table=11

/product="50S ribosomal protein L32"

/protein_id="AAM80423.1"

/db_xref="GI:21905568"

/translation="MAYPARHTSKAKNKRTHYKLTAPSVQFDETTGDSRSRHSVL
KGYKGRKIAKANEA"

gene

1898..2047

/gene="rpmg"

/note="synonym: SpyM3_1817"

1898..2047

/gene="rpmg"

/note="best blastp hit: gb|AAK34796.1| (AE006634) 50S
ribosomal protein L33 [Streptococcus pyogenes M1 GAS], and
gb|AAL98636.1| (AE010120) 50S ribosomal protein L33
[Streptococcus pyogenes MGAS8232]"

/codon_start=1

/transl_table=11

/product="50S ribosomal protein L33"

/protein_id="AAM80424.1"

/db_xref="GI:21905569"

/translation="MRVNITLHESGERLYTSKNKRNTPDRQLKKYSPKLRKHT
FTEVK"

gene

2328..2942

/gene="cadD"

/note="synonym: SpyM3_1818"

2328..2942

/gene="cadD"

/note="best blastp hit: gb|AAK34797.1| (AE006634) putative
cadmium resistance protein [Streptococcus pyogenes M1

GAS]"

/codon_start=1

/transl_table=11

/product="putative cadmium resistance protein"

/protein_id="AAM80425.1"

/db_xref="GI:21905570"

/translation="MIONVVTSLIYSSTAVDLIIILMFFAKRKRODIINILVGF
LGSVSLILSLFAFVLDYIPSKELGLGLPIPLGLKVLILLGDSGGAIAKEGLSK
DNKNLILVAMITFASCGADNIGVPVYFTTLNLANLIVALLTLVMIYLLVFSQAQL
AQVPSVGETLEKYSRWFIAVVLGLMYILIIENNSFDMIAVLG"

2954..3292

/gene="cadX"

/note="synonym: SpyM3_1819"

2954..3292

/gene="cadX"

/note="best blastp hit: gb|AAK34798.1| (AE006634) putative
cadmium efflux system accessory protein [Streptococcus

pyogenes M1 GAS]"

/codon_start=1

/transl_table=11

/product="putative cadmium efflux system accessory

protein"

/protein_id="AAM80426.1"

/db_xref="GI:21905571"

/translation="MKKDSICQGVINOQNVITATNYLEKEKVKOKSLRILSKFTDNKO
INIIFLLAVEELCVCDIACLLNLSMASASHLRLKLANQNILDRREGKIIYFIKDE
EIRDFNQLG"

3343..4266

/gene="SpyM3_1820"

3343..4266

/gene="SpyM3_1820"

/note="best blastp hit: gb|AAK34799.1| (AE006634)

hypothetical protein [Streptococcus pyogenes M1 GAS]"

/codon_start=1

/transl_table=11

/product="hypothetical protein"

/protein_id="AAM80427.1"

/db_xref="GI:21905572"

/translation="MKFFENYSOEIPIRIKLNKKHNITOSELGNAGOVSOVESGKR
PITSSMLVNLALTASSYIVIVFGEDEFIENLHVFFSSILYVDELEAVDENLVSFMS
DLSISQSSCLSIKTFANFNQKRWISITETEMDTFHKDDIDVWVGSEYNPARS
FRNNPNIILVIDFEEMADILLMLRNLIRSFENVCNLTFLDKXNGEFTFNLDKI
DSINKMWSNVSFTIIPNLIKRNPENFNIGFMYNDILERMKYKEDIPKSYLTSVPL
4332..5084

/gene="SpyM3_1821"

4332..5084

/gene="SpyM3_1821"

/note="best non-GAS blastp hit: gb|AAK79835.1|AE007695_8

(AE007695) FtsK-like DNA segregation ATPase, YDCQ

B. subtilis ortholog [Clostridium acetobutylicum]"

/codon_start=1

/transl_table=11

/product="hypothetical protein"

/protein_id="AAM80428.1"

/db_xref="GI:21905573"

/translation="NAIPDLKFSNLSQGIKPIVKHFLNLDKSHWIAIGNSGSK
PYATYFLSVLKMWSGLIIPDKFTPSQWARENKIAVHVENHKSDFVSQVNEOL
NOCATLIGKQALIDNPNHOFTHLTIVIDEVLALSEGKNKKEAFPSLSQIALLG
HATKHLIGKSQRFDHTIPISVREQNLVLLQIGNINOKTQFLFPDLPDEGIVIPWG
HGTGLIQVDNEHSYQVLLPCLLPYTKRGIL"

5081..5683

/gene="SpyM3_1822"

5081..5683

/gene="SpyM3_1822"

/note="best blastp hit: gb|AAK34801.1| (AE006634)

hypothetical protein [Streptococcus pyogenes M1 GAS]"

/codon_start=1

/transl_table=11

/product="hypothetical protein"

/protein_id="AAM80429.1"

/db_xref="GI:21905574"

/translation="MTTKTYPMRSFNLIYVLLLIAPICVYFGWLFNLVSVQLSTVTFV


```

NLDVKNVNTDISSVCYGFALISLLALLSLKMIORLNTDSIGNYCOSIYRTFLHR
FIPOREYQVKSNEEHOTVTTINPHTNGFNRAVRKICIVDRKDAITVELKVPDRQCG
KILKDTGVQLKEEIVIGHDYFESSPVVRNQLWFGQR"
complement(6159..6323)
/gene="SpyM3_1823"
complement(6159..6323)
/gene="SpyM3_1823"
/note="best blast hit: gb|AAK34556.1| (AE006610)
hypothetical protein [Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA080430.1"
/db_xref="GI:21905575"
/transl_table="MLFGTLLIAISYPIINSLWGWGYPVAILWLIIMRSALRIERE
KTYDQVQTP"
complement(6651..6809)
/gene="SpyM3_1824"
complement(6651..6809)
/gene="SpyM3_1824"
/note="best blast hit: emb|CAB65345.1| (Y18528) PR7
protein [Plasmodium falciparum]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AA080431.1"
/db_xref="GI:21905576"
/transl_table="MFLINNTNFEDIFLKHVFIKSLIVKVVNVIYKDKCPYDNY
QKDKRKA"
6894..7736
/gene="SpyM3_1825"
6894..7736
/gene="SpyM3_1825"
/note="best non-GAS blast hit: gb|AAK99693.1| (AE008462)

Query Match          47.7%; Score 598.2; DB 1; Length 53896;
Best Local Similarity 58.3%; Pred. No. 2.1e-141;
Matches 844; Conservative 0; Mismatches 386; Indels 3; Gaps 1;

QY 10 TTAATAAACCTCATACTGTGGCTTTAGTATTTTGGGTACTGTGTATTACGTC 69
DB 33882 TTTAAATAAACTTTAATGTGTTTATCTTTTGTGATATCTATCTGATTTACTA 33941
QY 70 AATGTTTATCTCTTGGTCTAAAGAGCTGTGCAATTAATAGCTTTTGTCTATAGCT 129
DB 33942 AATATGTATCTTTTTGG---AATCACTAGTGTAGGAATTTATGGAGTAATATAAACC 33998
QY 130 TACCTATTAGTCAAAATGTCCTTATCTTTTACAGCCATTTAAGGAGGCTGGG 189
DB 33999 TATCTAGTTATTAACCTGGATTAATCTTCCCTTAGGCCATTTAAGGAATCCACAT 34058
QY 190 CAATATAAGTTGACGCCATTATTCCTCTTATAACGAAGATGCTGAGTCAATGCTAGAG 249
DB 34059 GACATATAAGTTGCTGTGAATTCCTTCTTATAATGAAGATCCGAGTCAATATTAGAA 34118
QY 250 ACCTTAAAGTGTTCACAGCAAACTATCCCTAGCAGAAATTTATGTTGTTCAGCAT 309
DB 34119 ACCTTAAAGTGTGTTAGCAGACACTATCCGTTATCAGAAATTTATATGTTGATGAT 34178
QY 310 GGAAGTCTGATGACAGAGTATTAAAGCGCAATGAGACTATGTCGTCACACTGGTGAC 369
DB 34179 GGGAGTTCAACACAGATGCAATACAAATTAATGAGAGTATGTAATAGAGAGTGGAT 34238
QY 370 CTATCAAGCAATGTCATTTGTCATCGGTGACAGAAAAATCAAGAAAGCGTCATGCACAG 429
DB 34239 ATTTGTGCAAGCGTTATCGTTCACCGTTCCTTGTGCAATAAAGAAAGCGCATGCTCAA 34298
QY 430 GCGTGGCCCTTTGAAGATCAGACGCTGATGCTTTTGTGACCGTTGACTCAGATCTAT 489
DB 34299 GCGTGGGCAATTTGAAGATCTGACGCTGACGTTTTTTTTTAACCGTAGATCAGATCTAT 34358
QY 490 ATCTACCCCTGATGCTTTTAGGAGGAGTTGTTTAAACCTTTAATGACCCAACTGTTTGGCT 549

```

```

DB 34359 ATCTATCCAAATGCCTTAGAGAACTCTCTAAAGGCTTCAATGATGAGACAGTTTATGCT 34418
QY 550 GCGACGGGTACCTTTAATGTCTAGAAATAGACAAACCAATCTCTTTAAACACGCTTGACAGAT 609
DB 34419 GCAACAGGACATTTGAATGCTAGAAACAGACAACTAATCTATTAAAGGACTTACAGAT 34478
QY 610 ATTCCGCTATGATAATGCTTTTGGCGCTTGAAGGAGCTGCCCAATCCGTTACAGGTAATATC 669
DB 34479 ATCCGTTAGATAATGCTTTTGGGGTGGAGCGTGTGCTCAATCATTAACAGGTAATAT 34538
QY 670 CTGTTTGTCTCAGTCCGCTTAGCGTTTACAGACGCGAGGTGGTTGTTCTTCAACATAGAT 729
DB 34539 TTAGTTTGTCTCAGACCATTTAGTATTATTATCGACGTAAGTGAATTTCTTCACTAGAG 34598
QY 730 AGATACATCAACAGACCTTCTGGGTATTCCTTGTAAGTATTGGTATGATGACAGGTCGTTG 789
DB 34599 CGCTATAAAATCAACAGATCTCTGTTTACCTGTTAGCAITGGGGATGATCGATGTTTA 34658
QY 790 ACCAACTATGCAACTGATTTAGGAAAGACTGTTTATCAATCCACTGCTAAATGTAATACA 849
DB 34659 ACAAAATATGCTATTGATTTAGAGCGCACTGTCTACCAATCAAGCTAGATGATGATCT 34718
QY 850 GATGTTCTGACAAAGATGCTACTTACTTGAAGCAGCAAAACCGCTGGAACAGTCCTTC 909
DB 34719 GATGTAACCTTTCCAAATTAATAAGTATTATTAAAGCAAAATTCGATGGAATAATCTTTT 34778
QY 910 TTTAGAGAGTCAATATTTCTGTTAAGAAATCATCAACAATCTTTTGTAGCCCTATGG 969
DB 34779 TTTAGAGAATCTATTATTTCTGTCAAAAATCTTTCTTAATCCCATCGTGTGCTTATGG 34838
QY 970 ACCATCTTGAAGTGTCTATGTTTATGATGCTGTTGTTTATTTCTGTGGTGGATTTCTTTGTA 1029
DB 34839 ACTATTTTCGAAGTCGTATGTTTATGATGCTGATTTGCGCAATTTGGAATCTTTTGT 34898
QY 1030 GGCATATGCAGAAATTTGATTTGGCTCAGGGTTTATAGCTTCTTCTGGTGATFATCTCAAT 1089
DB 34899 AATCAAGCTATTCAATTAGACCTTATTAACCTTTTGGCTTTTATCCATCATCTTATC 34958
QY 1090 GTTGCCGCTGTGCGAACATTCATTACATGCTTAAGCAGCCGCTGCTCTCTGTTATCT 1149
DB 34959 GTTGCTTTATGCTGTAATGTTTCATTATATGCTCAAACTCTGCTAGTTTGTGTTATCT 35018
QY 1150 CGCTTTTATGGGTGCTGCTATGTTTGTCTACAGCCCTTGAATATATATCTCTCTTTT 1209
DB 35019 CCTCTGTATGGAATTAATACACTTTGTTGCTTACAGCCCTAAACTTTATCTTTATGC 35078
QY 1210 ACTATTAGAAATGCTGACTGGGACACAGTAAAAA 1244
DB 35079 ACCATTAATAATAGGAATGGGAATGGGGAACACGTAAAAA 35113

```

```

RESULT 12
AP005146 323825 bp DNA linear BCT 13-JUN-2003
LOCUS Streptococcus pyogenes SSI-1 DNA, complete genome, section 6/6.
DEFINITION AP005146 BA000034
ACCESSION AP005146.1 GI:28811737
VERSION
KEYWORDS
SOURCE Streptococcus pyogenes SSI-1
ORGANISM Streptococcus pyogenes SSI-1
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE
1
AUTHORS Nakagawa,I., Kurokawa,K., Yamashita,A., Nakata,M., Tomiyasu,Y.,
Okahashi,N., Kawabata,S., Yamazaki,K., Shiba,T., Yasunaga,T.,
Hayashi,H., Hattori,M. and Hamada,S.
TITLE Genome Sequence of an M3 Strain of Streptococcus pyogenes Reveals a
Large-Scale Genomic Rearrangement in Invasive Strains and New
Insights into Phage Evolution
JOURNAL Genome Res. 13 (6), 1042-1055 (2003)
MEDLINE 22683278
REFERENCE 2 (bases 1 to 323825)
AUTHORS Yamashita,A., Nakagawa,I., Kurokawa,K., Nakata,M., Tomiyasu,Y.,

```


Yamazaki,K., Okahashi,N., Kawabata,S., Yasunaga,T., Hattori,M.,
 Hayashi,H. and Hamada,S.
 Direct Submission
 Submitted (01-MAY-2002) Ken Kurokawa, Osaka University, Genome
 Information Research Center; 3-1, Yamadaoka, Suita, Osaka 565-0871,
 Japan (E-mail:Ken@gen-info.osaka-u.ac.jp,
 URL:http://www.gen-info.osaka-u.ac.jp/, Tel:81-6-6879-8365,
 Fax:81-6-6879-2047)
 genome project
 This clone was isolated from a patient presenting with toxic shock
 like syndrome.

FEATURES
 source
 1. 323825
 /organism="Streptococcus pyogenes SSI-1"
 /mol_type="genomic DNA"
 /strain="SSI-1"
 /db_xref="taxon:193567"
 complement(422..1012)
 /gene="SPs1572"
 complement(422..1012)
 /gene="SPs1572"
 /note="similar to GB:AAK33432.1 (AE006502) percent
 identity 100 in 196 aa"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="putative ATP-dependent protease proteolytic
 subunit"
 /protein_id="BAC64667.1"
 /db_xref="GI:28811738"
 /translation="MIPVIEQTSKRSYDIYSRLKKRIIMLTGPVEDNMSNVIA
 QLFLDAQNTKDIYLVNTPPGSVAGLAIVDMNFIKADYQTIIVMGMAASCTVIA
 SSGTKGRMLPNAEYMHGPMGGTGGTQTDNAIAAEHLKTRHREKLEIAQAGK
 TIKQIKDAERDYNMSAEETLAVGFIDEMENELK"
 complement(1235..1864)
 /gene="SPs1573"
 complement(1235..1864)
 /gene="SPs1573"
 /note="similar to GB:AAK33430.1 (AE006502) percent
 identity 100 in 209 aa"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="putative uracil phosphoribosyltransferase"
 /protein_id="BAC64668.1"
 /db_xref="GI:28811739"
 /translation="MGKQVISHPLQHKLSILRRQTSFDPRELVEIAMLGVGY
 SRDLPEVDIQTVPSTYKQKLAGKLAIVPLRAGIGVWDGLSLVPAKVGHIGM
 YRNEETLEPEYLVKLPEDINQRIPLVDPMLATGGSAILAVDSLKRGAANIKFVCL
 VAAPEGVKLOEAPHDIDIFTAALDDHLNHERGYIVPGLGDAGDRLEFGTK"
 complement(2032..3666)
 /gene="SPs1574"
 complement(2032..3666)
 /gene="SPs1574"
 /note="similar to GB:AAK33429.1 (AE006501) percent
 identity 99 in 544 aa"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC64669.1"
 /db_xref="GI:28811740"
 /translation="MSTKKQLTFQELMVGGAAMSTAGNFISRLGLVLIYIPWYIMWG
 QYAIQANALFNGYVAFYLLISTTGLNVAIAKQAVKNSMGQTEHVSQILSTIKL
 MLGLIFSAIMYLGSPFLASISGGDDTVPIMHLSLAVFIPVMSVIRICIQGHNN
 IKPVATSOAEQLRVITWMLTTFILMKLGSDYASAVTQSTFPAFICNVMAGVLGY
 YLWKQSLAAFPSSDIDIKLLETLESIPFIVTGSIAQAFQIDQMTFVNT
 LMTQDYSQQLLVFGYFNANPKTIVYLAASIGGVGIALITENYKDKMAAA
 RLININEMLVNFIIPALTGAILARPLYSVFYGASDRAIHLFVAVLFOTLLMALYT
 LFSPLQALFNKRAIYVFAYGILKVLQIPLIYLHAYGPLLATTALVPIYLMY
 RRLYQVTFNRKLIKQRLITLITELLMGLVYFVANWLLGYAFKPTGRLSLYLILII
 GGLGMVYATLTLTHOLDKLGSKASRLRQKGNH"
 3734..5199

gene
 CDS

CDS
 /gene="SPs1575"
 3754..5199
 /gene="SPs1575"
 /note="similar to GB:AAK33428.1 (AE006501) percent
 identity 99 in 481 aa"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="putative UDP-N-acetylmuramoylalanyl-D-glutamyl-2
 ,6-diaminopimelate ligase"
 /protein_id="BAC64670.1"
 /db_xref="GI:28811741"
 /translation="MITIQELLDILKKDNHREVLDADEYHYHQGFSPERLSYDSRQ
 VDGKTLFFAGATFKADYIKKEATLNGQLYISEVDYELGIVPLVTDIKKMSLIFAMA
 FYGNPQKIKLAFATGKTGTAAYFAYHMLKESYKPAFMSTMTTLDGTFKFSQLT
 TPESLDLFAAMAEVNTNGMTHLIMEVSSQAYLVDRVYGLTFDVGVLNLSPOHIGPIE
 HPTFEDYFHKRLLMNSRAVINSVMDFSLADQVADQEHFYGLPSDNOITTSOA
 FSTFAKQLAGHYDIQLIGHENQENNAAGLACLRGLASLADIKQGIATKTRYPGRMEV
 LTMTHAKVFVDYAHNGDSLEKLLSVVEHQTKMLLILGAPNGKESRRADFGRVH
 QHPNLVILTADDPNFEDPDISKEITASHIARPELIIISDREQALOKAMSLCOGAKDAV
 IIAAGKADAYQIVKQQVAYAGDLAIKHYL"
 5445..6227
 /gene="SPs1576"
 5445..6227
 /gene="SPs1576"
 /note="similar to GB:AAK33427.1 (AE006501) percent
 identity 99 in 260 aa"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="putative ferrichrome ABC transporter
 (ATP-binding protein)"
 /protein_id="BAC64671.1"
 /db_xref="GI:28811742"
 /translation="MTTISAEEDLAIYEQRTIIDLKSLFYIPEKTIITIGANGCKSKS
 LLKALTRLLPPKGVYNGQVNIATLEKEVAKLALPQVQATRTIVVLYELSYGR
 PFHOSYGNLSPADKKAITHMAQNTVMYAVQDPVADLSSGQRCQRYLAMAALAGTGT
 IFDEPTVYLDLNLHOLEILELVKSLNKDAGKTIIMVLDLNLNLSARYSDHLIMKHGKI
 FYGTGTDVMTSPITQIDIFQIKPVIYVDDPIHNCPIVLYTQLQ"
 6253..7185
 /gene="SPs1577"
 6253..7185
 /gene="SPs1577"
 /note="similar to GB:AAU97176.1 (AE009984) percent
 identity 99 in 310 aa"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="putative ferrichrome ABC transporter
 (ferrichrome-binding protein)"
 /protein_id="BAC64672.1"
 /db_xref="GI:28811743"
 /translation="MKKTLTLLTCLTITILACGNQATNHSNTASKSLSPMPQIAGY
 TTYGDIKPKQKRVVSLASTYTGKLLKLDMLNMGVTSYDKKNPILAKTVKKAKQAVTD
 LEATITKPDLLIVGSTEENIKQALAEIAPVISIEYKRKYDLVLSDFGRIFNKQKAK
 KWLKDWKTAAEYKEVKAATGKATFTIMGLYKEDVYLFKQDGRGGGFIHQAFHYD
 APEKVKTEFKQGLSLSQEVLPDYIGDVVIAAEDDKTGSALYESKLWQSIQPAVKH
 HVIKVANVSYFFDPLSLYQLETLREALISSEN"
 7196..8227
 /gene="SPs1578"
 7196..8227
 /gene="SPs1578"
 /note="similar to GB:AAU97175.1 (AE009984) percent
 identity 99 in 343 aa"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="putative ferrichrome ABC transporter (permease)"
 /protein_id="BAC64673.1"
 /db_xref="GI:28811744"
 /translation="MCLERFCLKKTQITKSHFWLFWLLSFLCVATYCHLRFQAV
 ALSHQDLNLSIFGKNGKHNVLLATRLPRLFGATLGSALAVSGTGMATIRPPIAE
 PGLLGINAGALALVLAFAVPHLHYSLLIILLSSGSLAATLVFGLSYQSGKGYHQL

gene
 CDS

RLVLAGAMVILLSALGOGITNYHLANAVIGWAGGLGVNWMQIGYIAPILILSIC
 LAQLLSHLTWLSSESQAKALGOKNLISAVFMILYLLLSRAVAIAGSIFGLVI
 PHLMKFTPHRYRLLPLCAVSGSFWVDIACRNLPPIETPLGALVSLIGFPCFL
 WLIIRGGRY*
 /gene . . . 9225
 /gene . . . SPa1579*
 CDS
 8188 . . . 9225
 /gene . . . SPa1579*
 /note= similar to GB:AAK33424.1 (AB006501) percent
 identity 98 in 345 aa
 /codon_start=1
 /evidence-not_experimental
 /transl_table=11
 /product="putative ferrichrome ABC transporter (permease)"
 /protein_id="BAC64674.1"
 /db_xref="GI:28811745"
 /translation="MLFVWNRKRRPLKSHRHLLGLYSLILLVLSLMGLALSLGESH
 LSFLLHVLGKSHSAISFVINRPLRIACLGGLGSLALLQLRITRNPGLADS
 GYLGTIGAGISLAIYVLSFSEFEQAHSHYPLFAMLGAVVTFVYVLSLKQGOID
 PYRLITGVAVTWSLMLVALVGHINRYKVDLVNWSGOLIGDWPVLSLVNLSFI
 CFMLITVSOAHFNLINGLADNTAIGLGLPLNRKRLILAVLAAGLALSLLVLSFI
 GLIAGHSTVLYVGSNKLITIPISILIGMILLVADTVGRVYLVGSNIQGLVSLIGA
 PYFLYMAKTK"
 complement(9270 . . . 9923)
 /gene . . . SPa1580*
 CDS
 complement(9270 . . . 9923)
 /gene . . . SPs1580*
 /note= similar to GB:AAI97173.1 (AB009984) percent
 identity 98 in 217 aa
 /codon_start=1
 /evidence-not_experimental
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="BAC64675.1"
 47.78; Score 598.2; DB 1; Length 323825;
 Best Local Similarity 68.36; Pred. No. 2.2e-141;
 Matches 844; Conservative 0; Mismatches 388; Indels 3; Gaps 1;
 QY 10 TTAATAAAGCTCAACTGTTGGGCTTTAGTATTTTGGGTACTGTTGATTTACGTC 69
 Db 303842 TTTAAAAAACTTAAATGTTTATCCCTTTATTTGTTGATATATCTGATTTATCTA 303901
 QY 70 AATGTTATCTTTGGTGTCAAGAAGCTTGTCAATTTATGCTTTTGGTGTGATGCT 129
 Db 303902 AATATGATCTATTTGG---ACATCAACTGTAGGAATTTATGGAGTAATATTAAATACC 303958
 QY 130 TACCTATTAGTCAAAATGCTTATCCCTTTTACAGCCATTTTAAAGGAGGCTGG 189
 Db 303959 TATCTAGTTTAACTTGATTTCTTCCCTTTATGAGCCATTTTAAAGGAATCCACAT 304018
 QY 190 CAATATAAGTTGAGCCATTTATCCCTTTTATAAGAGATGCTGATCATTTGCTAGAG 249
 Db 304019 GACTATAAAGTTGCTGTGTAATCTTCTTATAAGAGATGCTGATCATTTATAGAA 304078
 QY 250 ACCTTAAAGTTGCTAGAGCAACCTATCCCTAGCAGAAATTTATGTTGATGAGAT 309
 Db 304079 ACCTTAAAGTTGCTAGAGCAACCTATCCCTAGCAGAAATTTATGTTGATGAT 304138
 QY 310 GGAAGTGTGTAGAGAGAGTATTAAGGCGCATTTAGAGACTATGTCGCTGACACTGTGAC 369
 Db 304139 GGGAGTTCAACAGATGCAATCAATTTATGAGAGTATGTTAAATAGAGAGTGGAT 304198
 QY 370 CTATCAAGCAATGCTATGTTTCATGGGTGAGAGAAATCAAGAGAAAGCTGATGACAG 429
 Db 304199 ATTTGTGGAAGCTATGCTGTCACCGTTCCTTGTCAATAAAGGAAAGCCATGCTCAA 304258
 QY 430 GCCTGGCCCTTTGAAGATCAGACGCTGATGCTTTTACCGTTGACTGACTGATCTAT 489
 Db 304259 GCCTGGCCCTTTGAAGATCAGACGCTGATGCTTTTACCGTTGACTGACTGATCTAT 304318
 QY 490 ATCTACCTGATGCTTTAGAGAGTGTGTTAAACCTTTAATGACCCCACTGTTTGTCT 549
 Db 490 ATCTACCTGATGCTTTAGAGAGTGTGTTAAACCTTTAATGACCCCACTGTTTGTCT 549

Db 304319 ATCTATCAATAGCCTTAGAGAACTCCTAAAAAGCTTCAATATGATGAGACGCTTTATGCT 304378
 QY 550 GCGACGGGTGACCTTAACTCAGAAATAGACAAACCAATCTCTTTAAACAGCTTTGACAGAT 609
 Db 304379 GCAACAGGACATTTGAATGCTAGAAACAGACAACTATCTATTACGCGACTTACAGAT 304438
 QY 610 ATTCGCTATGATTAATGCTTTTGGGCTTGAACAGAGTGCCTCAATCCGTTACAGGTAATATC 669
 Db 304439 ATTCGCTATGATTAATGCTTTTGGGCTTGAACAGAGTGCCTCAATCCGTTACAGGTAATATC 304498
 QY 670 CTGTGTTGCTCAGTCCGCTTAGGCTTTACAGACGAGGTGGTGTCTCCTCAACATAGAT 729
 Db 304499 TTAGTTGCTCAGACCATTTGATTTATTCGAGCTGAAGTGAATATCTCCTAACCTAGAG 304558
 QY 730 AGATACATCAACAGACCTTCTGCTGCTTCTGCTGATTTGCTGATTTGCTGATGACAGTGTGTTG 789
 Db 304559 CGCTATAAAATCAACATCTCTAGTCTTACCTGTTAGCTTGGGATGATGATGATGTTA 304618
 QY 790 ACCAATCATCACTGATTTAGGAGAGCTGTTTATCAATCCACTGCTAAATGATTTACA 849
 Db 304619 ACAATATCTGCTATGATTTAGGAGAGCTGTTTATCAATCCACTGCTAAATGATTTACA 304678
 QY 850 GATGTTCTGACAGATGCTTACTTACTTGAAGCAGCAAAACCGCTGGAAAGTCTCTTC 909
 Db 304679 GATGTACCTTTCCAAATTTAAAGTATTTTAAAGCAACAAATCGATGGATTAATCTTTT 304738
 QY 910 TTTAGAGAGTCCATTTATTTCTTTGAAGAAATCAAGAAATCCCTTTTGTAGCCCTATGG 969
 Db 304739 TTTAGAGAGTCCATTTATTTCTTTGAAGAAATCAAGAAATCCCTTTTGTAGCCCTATGG 304798
 QY 970 ACCATATCTGAGGTGCTATGTTTATGATGCTGTTTATCTGCTGCTGATTTCTTTGTA 1029
 Db 304799 ACTATTTTCAAGTCTGATTTATGATGCTGTTTATCTGCTGCTGATTTCTTTGTA 1089
 QY 1030 GGCATGTCAGAAATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1089
 Db 304859 AATCAAGCTATTTCAATTTAGAGCTTTTAAACCTTTTGGCTTTTATCCATCATCTTTATC 304918
 QY 1090 GTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149
 Db 304919 GTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 304978
 QY 1150 CGGTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1209
 Db 304979 CCTCTGATGGAATATACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 305038
 QY 1210 ACTATTAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1244
 Db 305039 ACCATTTAAATACGGAATGGGGAACACGTAATAA 305073
 RESULT 13
 AF518732
 LOCUS
 DEFINITION Synthetic construct has operon deletion mutant, partial sequence.
 ACCESSION AF518732
 VERSION AF518732.1 GI:24415825
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 REFERENCE
 1 (bases 1 to 1893)
 AUTHORS Walker, J.A. and Timoney, J.F.
 TITLE Construction of a stable non-mucoid deletion mutant of the
 Streptococcus equi pinnae vaccine strain
 JOURNAL Vet. Microbiol. 89 (4), 311-321 (2002)
 MEDLINE 22272956
 PUBMED 12383640
 REFERENCE
 2 (bases 1 to 1893)
 AUTHORS Walker, J.A. and Timoney, J.F.
 TITLE Direct Submission
 JOURNAL Submitted (05-JUN-2002) Veterinary Science, University of Kentucky,
 108 Gluck Equine Research Center, Lexington, KY 40546, USA

FEATURES	Location/Qualifiers	Accession
source	1..1893	AF065404
gene	/organism="synthetic construct"	AF065404.1
gene	/mol_type="genomic DNA"	GI:4894216
gene	/db_xref="taxon:32630"	
gene	<1..>1893	
gene	/gene="has operon deletion mutant"	
gene	<1..>423	
gene	/gene="hasA"	
misc_feature	/note="hyaluronate synthase"	
gene	430..1137	
gene	/note="derived from vector pTW100"	
gene	<1138..1356	
gene	/gene="hasB"	
gene	/note="UDP-glucose dehydrogenase"	
gene	1406..>1893	
CDS	/gene="hasC"	
CDS	1406..>1893	
CDS	/gene="hasC"	
CDS	/codon_start=1	
CDS	/transl_table=11	
CDS	/product="UDP-glucose pyrophosphorylase"	
CDS	/protein_id="AAN59902.1"	
CDS	/db_xref="GI:24415826"	
CDS	/translation="MTKVKRAIIPAAAGLGRPLPATKALAKEMLPVIOKPTQIFVEE ALKAGIEILVVTGKARSIEDHDSNFEYNQAKGTELLKLVDETTAINLHFIR QSPHGRGLGDAVLAQAKFVGNFVWMLGDDLDITNFSAPLTKLIEDYDCTHASTI AV"	
BASE COUNT	559 a 427 c 385 g 522 t	
ORIGIN		
Query Match	32.6%; Score 409.2; DB 12; Length 1893;	
Best Local Similarity	92.9%; Pred. No. 2.7e-93;	
Matches	429; Conservative 0; Mismatches 33; Indels 0; Gaps 0;	
QY	475 GACTCAGACTATATCTACCTGATGCTTTAGAGAGTGTGTTAAACACCTTTAATGAC	534
Db		
QY	1 GACTCAGACTATATCTACCTGATGCTTTAGAGAGTGTGTTAAAGACCTTTAATGAC	60
QY	535 CCAACTGTTTCTCGACGGGACCTTAATGTCAGAAATAGACAAACCAATCTCTTA	594
Db		
QY	61 CCAACTGTTTCTCGACGGGACCTTAATGTCAGAAATAGACAAACCAATCTCTTA	120
QY	595 ACACGCTTGACAGATATTCGCTATGATAATGCTTTTGGCGTTGAACGAGTGCACCAATCC	654
Db		
QY	121 ACACGCTTGACAGATATTCGCTATGATAATGCTTTTGGCGTTGAACGAGTGCACCAATCA	180
QY	655 GTTACAGTAATATCTCTTGTTCACAGTCCGCTTACGTTTACAGACGCGAGTGTGTT	714
Db		
QY	181 GTTACAGTAATATCTCTTGTTCACAGTCCGCTTACGTTTACAGACGCGAGTGTGTT	240
QY	715 GTTCTACATAGATAGATACATCAACAGACCTTCTCGGTATTCCTGTAAGTATCGT	774
Db		
QY	241 GTTCTACATAGATAGATACATCAACAGACCTTCTCGGTATTCCTGTAAGTATCGT	300
QY	775 GATGACAGTGTGTTGACCACTATGCAACTGATTAGAAAGACTGTTTATCAATCCACT	834
Db		
QY	301 GATGACAGTGTGTTGACCACTATGCAACTGATTAGAAAGACTGTTTATCAATCCACT	360
QY	835 GCTAAATGATATACAGATGTTCCGACAGATCTCTACTTACCTGACGAGCAAAACCGC	894
Db		
QY	361 GCTAAATGATATACAGATGTTCCGACAGATCTCTACTTACCTGACGAGCAAAACCGC	420
QY	895 TGGAAAGACTCTCTTTAGAGAGTCCATATTATTCTGTTAAG	936
Db		
QY	421 TGGGATCCACTAGTCTAGAGCGCGCCGCCCGGTGGAG	462
RESULT 14		
AF065404	181654 bp DNA circular BCT 20-OCT-1999	
LOCUS		
DEFINITION	Bacillus anthracis virulence plasmid pX01, complete sequence.	

ACCESSION	AF065404	
VERSION	AF065404.1	GI:4894216
KEYWORDS		
SOURCE	Bacillus anthracis	
ORGANISM	Bacillus anthracis	
REFERENCE	1 (bases 1 to 181654)	
AUTHORS	Okinaka,R.T., Cloud,K., Hampton,O., Hoffmaster,A.R., Hill,K.K., Keim,P., Koehler,T.M., Lamke,G., Kumano,S., Mahillon,J., Manter,D., Martinez,Y., Riche,D., Svensson,R. and Jackson,P.J.	
TITLE	Sequence and organization of pX01, the large Bacillus anthracis plasmid harboring the anthrax toxin genes	
JOURNAL	J. Bacteriol. 181 (20), 6509-6515 (1999)	
MEDLINE	99445483	
PUBMED	10515943	
REFERENCE	2 (bases 1 to 181654)	
AUTHORS	Okinaka,R.T., Cloud,K., Hampton,O., Hoffmaster,A., Hill,K., Keim,P., Koehler,T., Kumano,S., Lamke,G., Manter,D., Martinez,Y., Riche,D.O., Svensson,R. and Jackson,P.	
TITLE	Direct Submission	
JOURNAL	Submitted (14-MAY-1998) Life Sciences Division, Los Alamos National Laboratory, TA43, JS-6, HRL-1, MS M888, Los Alamos, NM 87545, USA	
FEATURES	Location/Qualifiers	
Source	1..181654	
	/organism="Bacillus anthracis"	
	/mol_type="other DNA"	
	/strain="Sterne"	
	/db_xref="taxon:1392"	
	/plasmid="virulence plasmid pX01"	
	complement(1127..1630)	
	/note="Hypothetical protein, Methanobacterium thermoautotrophicum, (AE000831)"	
	/codon_start=1	
	/transl_table=11	
	/product="pX01-01"	
	/protein_id="AAD32305.1"	
	/db_xref="GI:4894217"	
	/translation="MEVLFELILIAVLIPNSVKKVKPKGKAGEKLVKRLMSKL DKPSYVHNVTYVEYDITQIDHIVIAETGVVETKNTYEWIGSEKAAKWTQGI FKKSFQNFPHONYKHAKIAEWLEIQQPCISMAAFHPKGLKRVNVHSKEKHVKYL HNKFAQ"	
	complement(1639..1644)	
RBS	2751..2755	
RBS	2760..3338	
CDS	/codon_start=1	
CDS	/transl_table=11	
CDS	/product="pX01-02"	
CDS	/protein_id="AAD32306.1"	
CDS	/db_xref="GI:4894218"	
CDS	/translation="MLEHTILSLVQLVPLVYVSTGIILILELTDITWETWSEYGF AIFLLIFQVFNKASIIISNOHKWFSKLNIFLVSLVLFVIFASPTLLKTAFLQM TYLFTLLECLIDRLPHNLYRVKSEICDKKQKIERIIPNNEEGHTTKTRDEQ YFLTYVLSGLISPISEVEFEYERNLMGR"	
CDS	complement(3389..4024)	
CDS	/codon_start=1	
CDS	/transl_table=11	
CDS	/product="pX01-03"	
CDS	/protein_id="AAD32307.1"	
CDS	/db_xref="GI:4894219"	
CDS	/translation="MKREKMGVLCVLMILLVPLQAQAVKVKFNKLGNDYS SSFQGLNNKTKTEKDFSGSIIQIKGLVAPKPKVGNIPISFVLGVIVMVAL ITKGNQWATGSMIFSFISMSRMLSVYELFSSDVSFNPITFDLDFKALALF TPPMLLAGIRLRHIESIKQPEYRWANRVITGSCLLCLMLAVLAPWLFANM"	
CDS	complement(4024..4653)	
CDS	/codon_start=1	
CDS	/transl_table=11	
CDS	/product="pX01-04"	
CDS	/protein_id="AAD32308.1"	
CDS	/db_xref="GI:4894220"	
CDS	/translation="MNINSISNNVKEKLEQTVYRKQOVLRVSKSSNIIHAYNIFP FEFSEVLSNDELSTSVKVIDRLSKQEIKNHSFTMCEVEEDGVIDDE TIFNTYVKKPPTIANKVKLVCGLGRLDLSKLHITVLNNINRIETEFYENENY	

RBS
CDS
VAYEKEIANATMGAYTKANIYYAADLYIAAAGNEIKLSIKPKLAGDV
complement(4034. .4039)
complement(4660. .4665)
complement(4989. .5309)
/codon_start=1
/transl_table=11
/product="px01-05"
/protein_id="AAD32309.1"
/db_xref="GI:4894221"
/translation="MEILEVEDMYQOEVSIPPEPEVFPVQQRVADVMDVFWRT
TKVLEQEQQLFVIMNSNNRYWHIEGEILLPENQLLYMKLEVDSPSYITVLEWSNRQ
LGN"
complement(5318. .5323)
complement(5720. .6220)
/codon_start=1
/transl_table=11
/product="px01-06"
/protein_id="AAD32310.1"
/db_xref="GI:4894222"
/translation="MLGDTGDLTLLIQNIKSQNLKGSTTEWTKAASGESKRVLANI
GNYTPFSEIEIQTNNVEKLPCTMDVAFMKEAQFQKAVECIKKEQVPCRVKQNV
LEVDGEIEDITRALTLTNDKRLVKQDITFWSVIVEVDPSSKKKGKKNKIIDGQY
EFDLPG"
complement(6231. .6236)
complement(6544. .8352)
/note="reverse transcriptase, Is629, E. coli, (AB011549);
reverse transcriptase, Is629 (574 aa) E. coli (AB011549);
306/543 positive aa (56%)"
/codon_start=1
/transl_table=11
/product="px01-07"
/protein_id="AAD32311.1"
/db_xref="GI:4894223"
/translation="MNTVNMKSASVSPQVTOWNSTOMKEITNVKKLRIORIVRAEQNGN
KRVKRLQMLRSGNALLSKRTVQINQCKRTAGDGIITNPEDRVKLEHLIAGY
SVNIKAPRVKRAVTPKKNKKRPLGIVIKDRFQNVKNALPEQWECRPFESMTGF
RPSRSHDAMNLEKLSRGTNRNMFEGDQFGCQDNLNHEILSCIEGPPYSNAINQ
WLAGCINDKFTVETGTPQGGIISPLIANIHLGHEKELGVRVHFPRKDGAMLYPD
SLGIYADADFFVVCNSKEAEMSAQLQPYLDKRLKLAETKRVHITDGFDFLFG
NFPQYPTGEGQLFKPSNGSVKKAKEKISEIPKSHRGRSIGOLIRKLNVPVIGIANY
WSPVAKVIGYDIDSVYKVKHHLKHKHKKARWLNKKYFPHDHTGVSDQKLLTD
PDNHKQLIRMTWPIVRHVLIKYKNSPDDANLKYFAKDEKIFNRNNSKKLAK
KTYKCRVNSLSGEEPLESHNIVPVIGKGDYNDLLELLHSCQKHQHALLEWYGN
GKQLPKVQYLRKSKPIFNSKKAIVTMLNTPKFKY"
complement(8362. .8367)
complement(8956. .11541)
/codon_start=1
/transl_table=11
/product="px01-08"
/protein_id="AAD32312.1"
/db_xref="GI:4894224"
/translation="MLQELIKQDTFDBQTPAILQLETGTASHSAFCFAMAINHREI
QFVAGLNDSTLKSFRRAISMGTSLRYEGEGKRESHYTLDRKVNVIKSGOFFINTQ
TINKKAITATSKOLETKYIVAIDEAPMOVDFLMAPPYGLPILFEWAKPIYEMLT
KNLLOPLEVTDKNFESSLSIAQVVLKEKDCKEFLSEMRSSKCOFFQEGTGKINKI
ODLNEVLEYSPVMDVAYKLEDEPLHQPMEQELSHFDIKRPLFPVQAHVATGAAGA
LQVQGIHQGEMSGSSIMTAVDGYFHLTGKGRVTCVFPPTLTERKAKEEIRH
LIPDAVHLIKRTEDLIRHQSWIQTGRPKPPTFFVISTTHRGDSIKQMDPLPYQ
IALSKSKEEVQRYKNGYPCDCAKRLKRTSSIMVHQAGEQKVCQYKDFGSDLI
DKPTNKSVCADCNISWPKYKVAASKDWTYKYNLQVAILKGNKPKQKOLEN
RVKPTDAKQGRVARKVATVEYIRBKMKHFDALICDEVHELKSETAQGTALGSLKV
CKVAVAGTFLGGASDVYILRFLPQKAVSGFEISSEMMNETYIGNIETVYHD
GEVTSNARSAGQGLDTRKVPVGPISPYTFQFMDDTINVELKDVMPNVELVNPYD
LIVDMSEQKQAYEHMKSEKAIIEAAGTDPAGKLWLSYINTGNAYPDNNIHPNYI
LDAGKEFNWKNDFKDFEDQPKKEGLILRTSEISREPTVIYVCDQTSVKEKRO
VQPRKLQVVEQIPGAKVAULRTNTVTPPKRSEMLKEQVNSGVNVIICSLVRYGLDC
ATSR"
complement(11549. .11554)
complement(11613. .11840)
/codon_start=1
/transl_table=11
/product="px01-09"
/protein_id="AAD32313.1"

RBS
CDS
/db_xref="GI:4894225"
/translation="MSVLLIGISLLIWEFSNRMFRIALVLTVICLCAQLGAGTHLE
VLEQFSVPHGVGDMJSAAQSLPEWKRIFDF"
complement(11847. .11852)
complement(11856. .12947)
/codon_start=1
/transl_table=11
/product="px01-10"
/protein_id="AAD32314.1"
/db_xref="GI:4894226"
/translation="MRLGNDLAAQFYPTPLTECKHLVRLQLMBSERAYACDPCCGSG
TLISFADETKQAGCYHTGYVELDANRYQAKQEOLOVYVPSSESMISHNYPILF
LNPPYNTLETNAKSEKMEFNLKRAHLYLQEGSIYVPIPIPIRDFARDISITLAK
VEELGLRFGDENEEFEQFKOCVFIRKRTATNTDIFNDRFANFENNSDDIDFVRKH
VSTLAQWNEKOWTPELRHONKLIPTSRVDYKDAYGKSGEIASLKKRLNRNGY
LLQCKTAARAKMPIASQGLGLLITGVADGLLGGDTLHVRGSENYVEHSFEQL
EKTTEVYLQKRRAKFITATPTGVVVELV"
complement(12956. .12961)
complement(13209. .13562)
/codon_start=1
/transl_table=11
/product="px01-11"
/protein_id="AAD32315.1"
/db_xref="GI:4894227"
/translation="MEDNOKSYAAPVMSSSESRVNMRCGEAEIILPDCKEYFEP
HALNEKATOEIALXIEIQEMKVGYYVILWRVGGQSYSGTKIVAEPRKILRKUK
DYFDEVEVYEPK"
complement(13574. .13579)
complement(14226. .14231)
complement(14237. .14929)
/codon_start=1
/transl_table=11
/product="px01-12"
/protein_id="AAD32316.1"
/db_xref="GI:4894228"
/translation="MLKNILICISFISFISLFIPTQDITMTKIVFIENGIGISI
PTIVAMQOLKOTILKDFIDNEMKETTKEITNFVGLLKLOVIOGDLKRLSHYIN
RESMHGLYREFRDKWKNAYIPIINMVAEKLIIDVVASLLELELOPHLOKIVLSY
VFTAEVYLNALPSLGGSEKDKRVKLLSRNIQKNVILSAGFEKVEKVEEETYTE
FTFVLQKNRG"
complement(15040. .19002)
/note="erythrocyte invasion and possible binding protein;
Rhopty protein (2401 aa), Plasmodium yoelii (U36927);
520/1194 positive aa (43%)"
/codon_start=1
/transl_table=11
/product="px01-13"
/protein_id="AAD32317.1"
/db_xref="GI:4894229"
/translation="NAKFEIPDIKYEDIKAAKTGNLILFVGAGISMLGLPSWKG
YAOILSDALEKGVIDSLEYELKEEDPKKILITCKMLFEKELIPKSNELPFKPKN
BEYKAIYEKLYSNNAIITTYWDECLDVLITPQSDNEKVSNDLSEENKLTUSSET

Query Match 22.6%; Score 282.8; DB 1; Length 181654;
Best Local Similarity 57.8%; Pred. No. 5e-61;
Matches 503; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 383 TCATTGTCATCGCTCAGAGAAAATCAAGAAAGCGTCATGCACAGCGCTGGCCCTTG 442
DB 111615 TAATGTCATCGTTTACCTAAGAAATTCGGGAAAGACATGCTCAATTATGGGCTTTA 111674
QY 443 AAAGATCAGAGCGTGTCTTTTGGACCGTGTACTCAGATCTTATCTACCCCTGATG 502
DB 111675 AACGACAAACAGCAGATGCTATTGTTACCATTCATTACAGATGCTGATTTGCCAAATG 111734
QY 503 CTTTACAGAGTGTGTTAAAAACCTTTAATGACCAACTGTTTTTGTGCGAGCGGTCAAC 562
DB 111735 CTGTTAGAGGATTATGAACCCCTTATGATGAAAAGATTAATGCCACAACGTGTCACG 111794
QY 563 TTAATGTCAGAAATACAGAAACCAATCTCTTACACGGCTGCACAGATATTCGCTATGATA 622
DB 111795 TGAACATTCGTAATAGAAATGATATTTAATAACAAAACAAATTTGATATGCGTTATGACA 111854

```
QY 623 ATGCTTTGGCGTTGAAGAGCTGCCCAATCCGTTACAGTAATATCCTCTGTTGTCACAG 682
Db 111855 AAGCGTCCGTGGAGCGTGCACGACGTCGTACAGGAATGTTCTGTTGTAGTG 111914
QY 683 GTCCGCTTAGCTTTACAGAGCGAGGCTGTTGTTCCCTAACATAGATAGATACATCAACC 742
Db 111915 GCGCGTTAAGTGTCTATCTGACAGAGTAATAACTGAAATTTAGACATATTGGAAGTC 111974
QY 743 AGACCTTCCTGGGTATCCTGTAAGTATTGGTGTAGCAGAGCGTGTACCAACTATGCAA 802
Db 111975 AGATGTTCCCTGGTGGAGGTCGAGTTTGGAGATGATAGATGCTAACTAATTATGCTA 112034
QY 803 CTGATTTAGGAAGAGCTGTTATCAATCCACTGCTAAATGTTATTACAGATGTTCTCGACA 862
Db 112035 TTTTGAAGGGAACAACTTATCAATCCACTGCTGATGTTATTACTGATGCTCCAACTA 112094
QY 863 AGATGCTTACTTACTTGAAGCAGCAAAACCGCTGGAACAAGCTCTCTTTAGAGAGTCCA 922
Db 112095 CATTAACAACAATTTCTTAACAGCAACTAGCTTGGAAACAAGTCATTTTTAGAGAAAGTT 112154
QY 923 TTATTTCTGTTAAGAAATCATGAACAATCCTTTTGTAGCCCTATGACCATCTACTGAGG 982
Db 112155 TAATTCACCTGCACTGTTGTAAGAAACCAAAATGTTCTTGTGGAACAATTTTCGAAA 112214
QY 983 TGTCTATGTTTATCATGCTGTTGTTTATTCCTGCTGATTTCTTTGTAGGCAATGTCAGAG 1042
Db 112215 TATCGTTATGATTTTATTTGGCTTTCCCTACTTCTTAAGTATTATTCTCAGGCAAGTC 112274
QY 1043 AATTTGATGGCTCAGGCTTTAGCCTTTCTGCTGATTAATCTTCTGCTGCTGCTGCTG 1102
Db 112275 ATGTAGGGTTAAATTTGGCTGTTTATTTATTTGGTTATTTATTCATAGCTGATATGCTA 112334
QY 1103 GGAACATTCATATCATGCTTAAAGCACCGCTGCTCTTCTGTTATCTCTGCTGCTGCTG 1162
Db 112335 GAATGATTTTATCTATTAACAATCCCTTACTTCTTACTGCTGCTGCTGCTGCTG 112394
QY 1163 TCTGCTGATTTGTTGCTTACAGCCCTGAAATATATATCTCTTTTACTATTAGAAATG 1222
Db 112395 TTCTCCATGTTATGACACTTATAGCTATAGCTTTTATGCTTACTTACTAATTAATCTA 112454
QY 1223 CTGACTGGGACACGTAAGAAATTTAT 1252
Db 112455 ATGGTTGGGGAACAGCTTAATTAACAGTAAT 112484

RESULT 15
AE011190
LOCUS 181677 bp DNA circular BCT 14-JUN-2002
DEFINITION Bacillus anthracis str. A2012 plasmid px01, complete sequence.
ACCESSION AE011190
VERSION AE011190.1 GI:20520075
KEYWORDS
SOURCE
ORGANISM
Bacillus anthracis str. A2012
Bacillus anthracis str. A2012
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
REFERENCE
1 (bases 1 to 181677)
Read,T.D., Salzberg,S.L., Pop,M., Shumway,M., Umayam,L., Jiang,L., Holtzapfle,E., Busch,J.D., Smith,K.L., Schupp,J.M., Solomon,D., Kelm,P. and Fraser,C.M.
Comparative genome sequencing for discovery of novel polymorphisms in Bacillus anthracis
Science 296 (5575), 2028-2033 (2002)
PUBMED 12004073
REFERENCE
2 (bases 1 to 181677)
Read,T.D., Salzberg,S.L., Pop,M., Shumway,M., Umayam,L., Jiang,L., Holtzapfle,E., Busch,J.D., Smith,K.L., Schupp,J.M., Solomon,D., Kelm,P. and Fraser,C.M.
Direct Submission
Submitted (09-MAY-2002) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
FEATURES
location/Qualifiers
```

source

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

```
1. .181677
/organism="Bacillus anthracis str. A2012"
/mol_type="genomic DNA"
/strain="A2012"
/db_xref="taxon:191218"
/plasmid="px01"
/note="Florida strain"
462..554
/gene="BXA0001"
462..554
/gene="BXA0001"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAM25958.1"
/db_xref="GI:20520076"
/translation="MVARIHLIKFKVVIKSKKHSNECLFLWTK"
complement(996..1109)
/gene="BXA0002"
complement(996..1109)
/gene="BXA0002"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAM25959.1"
/db_xref="GI:20520077"
/translation="MLYGLFGLVISIVAIRLANKYKLGKGNKDFVKKNG"
complement(1127..1765)
/gene="BXA0003"
complement(1127..1765)
/gene="BXA0003"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAM25960.1"
/db_xref="GI:20520078"
/translation="MEVLFELILIAVLPLNSVYKHKVKKWKGAGEKLKVRMLSKL
DPKSYVHLNVTYVTEGDTQIDHIVIAETGVFVETKNYEGTYSSEKAARWTOGI
FRKSSFPNPHONYKHAKIEMLEIQLPFCISMAAPHPKCSLKRNVHSKSHVLYY
NDLOKCIESTVDVLNDEVOHIVHTILRANIMDXDIEKKHKVLYHNKFAQ"
complement(1782..2165)
/gene="BXA0004"
complement(1782..2165)
/gene="BXA0004"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAM25961.1"
/db_xref="GI:20520079"
/translation="MNRREYIALRNSISIGMLLTICIVIGLSYFSSVYLLIGCFIL
SIKTYGLMGTYKNTKRLPFWANFGVIEKPVNRVTRPFPQFGLLVSLALGFVEMVIK
GIKTSYIFLIGLTHVPRLMRRNRM"
2291..2875
/gene="BXA0005"
2291..2875
/gene="BXA0005"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAM25962.1"
/db_xref="GI:20520080"
/translation="MLKLVNNTTEPEGDSAFSNFTCKDCVYLSKSDCSLQAA
DSNPLPYPCGNFEYETIGTDLNSHIOQLREKAQILAEYSQAQPSIDPDKO
ALWGSVNYGCWLIINDSKRFALAEQVPQDNMSSKKYASFPPLHDHGAKEGRHYM
CWTYNEKGISEYHLLIGVIMKIAEGVKLGPNY"
2895..3473
/gene="BXA0006"
2895..3473
```

```

/gene="BXA0006"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein, (pX01-02)"
/protein_id="XAM25963.1"
/db_xref="GI:20520081"
/translation="MLEHTILSLVOLPLIVSTGIYILLELIDTEWTKWFYEGIF
AIFLTFQVONKKSIIISLNOHKFWSKINIFLVLFSIPITLLKIAFILOM
TVILFLELCYIDRLPIHNLVRYKKEICDKQLKIERIPNNEHGHTHITKREQ
YFIYTAVILGSLSPISEEVEEYNNLMGR"
complement(3524..4111)
/gene="BXA0007"
complement(3524..4111)
/gene="BXA0007"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein, (pX01-03)"
/protein_id="XAM25964.1"
/db_xref="GI:20520082"
/translation="MILLIPLQAAQASQVDFENKLGNDVSSFSKDLGNRTKITEK
DSFGSIIQIEGLYAPAFKGVNIFISFVLGVIVMVALITKNGQWKNATGSKI
FSFISMSRLSVFLPSSDFSNFIPTDPLDLKATALIPTPLIAGIRLRRIHE
STKQPEYRANRVITGSCILCMVLAVLAPLFAFM"
complement(4159..4788)
/gene="BXA0008"
complement(4159..4788)
/gene="BXA0008"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein, (pX01-04)"
/protein_id="XAM25965.1"
/db_xref="GI:20520083"
/translation="MNIINSNNVKELOTVREKKQVQLRVSKSSNHYMIFLP
EFSEYKINTELDEYINLSVSVYIDRLSQEIEKONHSEFIMTCEVEDVIDDE
THENTTVKPTTANKVKPLVGLGNELDLKHILHTVLNNINRIEIEFINENETI
VAVEKEELANAGTKANIYYAADLYAAEGNEIKLASINKPLAGDV"
complement(5124..5837)
/gene="BXA0009"
complement(5124..5837)
/gene="BXA0009"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein, (pX01-05)"
/protein_id="XAM25966.1"
/db_xref="GI:20520084"
/translation="MFTKIFIGFYVIVVCSNKKPLREKKKEMENKPLRPVY
WCPRENKPIGQERITVTKETLEKDVASIVDYVSVHKNLGEVSDRLKHEDIE
EVEYAEDEHPTPEAVYDIHFPSEDIPELLEYEDMFPQOEIVSIPPEPEVPVQ
QVAVYKMDERWTRIVLEQEGGLRVNNSNKKRYIHWTEGILPENOLLYMKLEY
DSPYSITLVSNRLOGNI"
complement(5854..6561)
/gene="BXA0010"
complement(5854..6561)
/gene="BXA0010"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein, (pX01-06)"
/protein_id="XAM25967.1"
/db_xref="GI:20520085"
/translation="MIFYQLPYSIFALNQAAHRHCRIGTKOCRTFYLVYRETFQOM
AQILAKKNAAEAMNGEATSGNLMLGDTGLQTLILIONIKSGNOLKSTEWKAA
SESRKLVANIVTIPVFESIEIOTNWKEVPTDMVAFMKAQFQVQAEVCKKE
QVPGVVRKQNVLEVDGEIEDIIRALTTLNTDKRMIVQKQDITFWDVSVIVEVPSKKK
KOKNKIIDQYEDPLRG"
complement(6678..8486)
/gene="BXA0011"
complement(6678..8486)
/gene="BXA0011"

```

```

/note="identified by match to PFAM protein family HMM
PF00078"
/codon_start=1
/transl_table=11
/product="reverse transcriptase/endonuclease protein,
(pX01-07)"
/protein_id="XAM25968.1"
/db_xref="GI:20520086"
/translation="MNTVMKSAVSPOVTOWNSIOWKIEINYVKLRQIRVRAEQFN
ARKVKQRLMRKANLLSIRKVTQINOGRKTAGIDGTTTTPEDRVKLFHLKGY
SVRNKAPVKRAYIPKNGKRPGLPIVKDRIFQNMVKNALFPOWECRESMSYGF
RPKSAHDAMANFLKLSRGTNRWIFEGDFQCFDNLNHEHILSCIEGFYSNAIQ
WLNAGCIDNKTKTFYKTTGTPOGTTISPLANIALHGMKEKELGVYHFKRGGAMLYPD
SIGIVRYADDVIVCNSKEEAESYAKLOPYLDKRLKLABEYRVVHTIDGDFDGF
NFRQYPTKEGSQLFKPSNQSVKAKEISEIFKSHGRSISGLIRLKNPVTIGIANY
PDNKNQILRMWTPIVRHVLIYKNSPDDANLKYFAKDERIKENRFNENSKSLAK
KTKYKRCVNNLSVGEESHNHVPKVIGGKDEYDNLLELLHCSCHKQHALLWYGN
GKQLPKVQAYLKSKIPINSKKAVRTMLNTFKPKY"
complement(8991..9083)
/gene="BXA0012"
complement(8991..9083)
/gene="BXA0012"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="XAM25969.1"

Query Match      22.6%; Score 282.8; DB 1; Length 181677;
Best Local Similarity 57.8%; Pred. No. 5e-61;
Matches 503; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

QY 383 TCATGTTTCATCGTCAGAGAAAATCAAGAAAGCGTCATGCACAGCGCTGGGCTTTG 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111666 TAATGTCATCGTTTACCTAGAATTCGCGGAAAGACATGCTCAATTATGGGTTTTA 111725
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 443 AAGATCAGACGCTGATGCTTTTTTGACCGCTGACTCAGATCTTATCTACCCGTGATG 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111726 AACGGACAACAGCAGATGCTATTGTTACCATGATTTCAGATGGTGATTGTTCCCAATG 111795
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 503 CTTTAGAGGAGTGTGTTAAACACCTTTAATGACCAACTGTTTTGTCGACGGGTGACC 562
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111786 CGTTTAGAGAGTTATTGAAACCTTTAATGATGAAAGAAATATGGCCACAACGTGTCACG 111845
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 563 TTAATGTCAGAAATAGACAAACCAATCTCTTAACCGCTTGACAGATATTCGCTATGATA 622
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111846 TGAACATTCGTATAGAATGATATTTATTAACAACAACTAATTGATATGCGTTATGACA 111905
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 623 ATGCTTTTGGCGTTGAACGAGTGGCCCAATCCGTTACAGGTAAATATCTTGTGCTCAG 682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111906 ATGCGTTTCGCTGGAGCGTCGACACAGTCCGTAACAGGAAATGTTCTTGTGTTAGTG 111965
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 683 GTCCCGCTTAGCGTTACACACCGCAGGTGGTTGTCCTAACATAGATAGATCATCAACC 742
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111966 GGCCTTAAGTGTGTATCGTAGAGAAGTAATAACTGAAATTTAGAACATTTAGGAAGTC 112025
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 743 AGACCTTCCTGGGTATTCTCTGTAAGTATTGGTGATGACAGGTGCTTGACCAACTATGCAA 802
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 112026 AGATGTTCTTGGTGAGAGGTGTCAGTTTGGAGATGATAGATGCTAATAATTATGCTA 112085
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 803 CTGATTTAGGAAGACTGTTTATCAATCCACTGCTAAATGTTATTACAGATGTTCTTGACA 862
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 112086 TTTTGAAGGGAAGAAACAGTTTATCAATCCACTGCTGATGTTATTACTGATGCTCAACTA 112145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 863 AGATGCTCTACTTACTTACAGCAGCAAAACCGCTGGAACAAGTCTCTTCTTTAGAGAGTCCA 922
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 112146 CATTAACAACATTTCTTAACAGCAACACTAGCTTGAACAACAGTCAATTTTITAGCAAGATT 112205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 923 TTATTTCTGTTAAGAAAATCATGAACAATCCTTTTGTAGCCCTATGACCATATCTTGGG 982
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 112206 TAATTCACITGGCATGGTATGAAAAAACCACAAATGTTCTTGTGTTGGACAATTTTCGAAA 112265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

QY	983	TGCTATGTTTATGAGCGCTGCTTTATTCGTGGTGGATTCCTTTGTGAGCAATGTCAGAG	1042
Db	112266	TATCGTTATGAGATTTTATTTGGGTTTCCCTACTCTAAGTATTTATCTCAAGCAAGTC	112325
QY	1043	AAATTGATTTGGCTCAGGGGTTTTAGCCTTTCTGGTGATTATCTTCATTTGTCGCCCTGTGC	1102
Db	112326	ATGTAGGGTTAAATTTGGCTGTTTATTTATTTGGGTTATTTTCATTACGCTGATATGCTA	112385
QY	1103	GGAACTCATTTATGAGCTTAAGCACCGCTGCTCCTCTGTTATCTCCGTTTTATGGGG	1162
Db	112386	GAAATGATTTTATCTATTAAACATCCGCTACTTCTTCTACTGGCGCAATTATATGAA	112445
QY	1163	TGCTGCTATTTGTTGGCTACAGCGCTTGAAATATATATCTCTTTTACTATTAGAAATG	1222
Db	112446	TTCTCCATGTTATGACACTATACCTATAGCGCTTTTATGCTTTTACTAACTATTAAATCTA	112505
QY	1223	CTGACTGTGGGACACGCTAAAAATTTATTAT	1252
Db	112506	ATGGTTGGGGAACACGTTAAATACAGTAAT	112535

Search completed: October 2, 2003, 10:30:04
Job time : 4792 secs

Result No.	Query			Length	DB	ID	Description
	Score	Match					
1	1254	100.0	1254	20	AA558841		Streptococcus equi
2	604.6	48.2	1512	25	AB57951		S. pyogenes hyalur
3	601.4	48.0	1335	24	ABN67167		Streptococcus poly
c	61.6	4.9	1653	24	ABL41013		Murine hyaluronan
5	61.6	4.9	1659	25	AB276735		Mouse hyaluronan s
6	61.6	4.9	2890	21	AA288200		Mouse hyaluronate
7	61.6	4.9	2947	19	AA118821		Murine hyaluronan
8	61.6	4.9	2948	24	AB199697		Mouse ischaemic co

XX PF 30-OCT-1998; 98WO-US23153.
 XX PR 26-OCT-1998; 98US-0178951.
 XX PR 31-OCT-1997; 97US-0064435.
 XX PA (OKLA) UNIV OKLAHOMA STATE.
 XX PI DeAngelis P, Kumari K, Weigel PH;
 XX WPI; 1999-337486/28.
 XX DR P-PSDB; AAY06206.
 XX PT Nucleic acid encoding hyaluronate synthase for production of
 PT hyaluronate acid with controlled molecular weight and targeting
 PT specificity
 XX PS Claim 3; Page 118; 125pp; English.
 CC This DNA sequence codes for a hyaluronate synthase, termed seHAS
 CC (see AAY06206), of group C Streptococcus equisimilis. To isolate
 CC the DNA, genomic DNA was subjected to PCR amplification using
 CC degenerate primers (see AAY58847-52) based on known related
 CC sequences, and PCR products were used as probes to screen a DNA
 CC library. Expression of the seHAS gene correlates with virulence
 CC of streptococcal Group A and Group C strains, by providing a means
 CC of escaping phagocytosis and immune surveillance. The invention
 CC also provides recombinant vectors containing the isolated DNA, and
 CC prokaryotic or eukaryotic host cells which produce seHAS and its
 CC hyaluronate acid product, particularly a product with modified
 CC structure or molecular size. The hyaluronate acid produced this way
 CC is purer than that produced by conventional methods. The seHAS
 CC gene also provides a new probe to assess the potential of bacterial
 CC specimens to produce hyaluronate acid.
 XX SQ Sequence 1254 BP; 335 A; 240 C; 261 G; 418 T; 0 other;

Query Match 100.0%; Score 1254; DB 20; Length 1254;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGACATTAATAAAGCTCATACGTTGGTGGCTTTAGTATTTTGGGTACTGTTG 60
 DB 1 ATGAGACATTAATAAAGCTCATACGTTGGTGGCTTTAGTATTTTGGGTACTGTTG 60
 QY 61 ATTACGTCATATCTCTTTGGTCTAAAGAAAGCTGTCAATTTATGGCTTTTG 120
 DB 61 ATTACGTCATATCTCTTTGGTCTAAAGAAAGCTGTCAATTTATGGCTTTTG 120
 QY 121 CTGATAGCTTACCTATTAGTCAAAATGTCCTTATCCCTTTTACAGCCATTAAAGGA 180
 DB 121 CTGATAGCTTACCTATTAGTCAAAATGTCCTTATCCCTTTTACAGCCATTAAAGGA 180
 QY 181 AGGCTGGGCAATATAAGTTGACGCAATTTCCCTTATACCAAGATGCTCACTCA 240
 DB 181 AGGCTGGGCAATATAAGTTGACGCAATTTCCCTTATACCAAGATGCTCACTCA 240
 QY 241 TTGCTAGACCTTAAAGTTGTTACGACGAAACCTATCCCTAGCAGAAATTTATGTT 300
 DB 241 TTGCTAGACCTTAAAGTTGTTACGACGAAACCTATCCCTAGCAGAAATTTATGTT 300
 QY 301 GTTACGATGAGTCTGATGACACAGCTATTAGCCATTAGCACTATGTCGGTGAC 360
 DB 301 GTTACGATGAGTCTGATGACACAGCTATTAGCCATTAGCACTATGTCGGTGAC 360
 QY 361 ACTGGTGACCTATACGAATGTCATTTGTCGCTCAGAGAAATCAAGGAAAGCGT 420
 DB 361 ACTGGTGACCTATACGAATGTCATTTGTCGCTCAGAGAAATCAAGGAAAGCGT 420
 QY 421 CATGCACAGCCCTGGGCTTTGAAAGATCAGACGCTGATGCTTTTGGCCCTTGACTCA 480
 DB 421 CATGCACAGCCCTGGGCTTTGAAAGATCAGACGCTGATGCTTTTGGCCCTTGACTCA 480

QY 481 GATACCTATTATCTACCTGATGCTTTAGAGAGTGTGTTAAAAACCTTTAATGACCAACT 540
 DB 481 GATACCTATTATCTACCTGATGCTTTAGAGAGTGTGTTAAAAACCTTTAATGACCAACT 540
 QY 541 GTTTTGTGCGAGCGGTCACTTAACTGTCAGAAATAGACAAACCAATCTCTTAACAGCG 600
 DB 541 GTTTTGTGCGAGCGGTCACTTAACTGTCAGAAATAGACAAACCAATCTCTTAACAGCG 600
 QY 601 TTGACAGATATTCGCTATGATAATGCTTTTGGGCTTGAACGAGCTGCCCAATCCGTTACA 660
 DB 601 TTGACAGATATTCGCTATGATAATGCTTTTGGGCTTGAACGAGCTGCCCAATCCGTTACA 660
 QY 661 GGTAAATATCCCTGTTTTCCTCAGTCCGCTTAGCGTTTACAGACGGGAGGTGTGTCT 720
 DB 661 GGTAAATATCCCTGTTTTCCTCAGTCCGCTTAGCGTTTACAGACGGGAGGTGTGTCT 720
 QY 721 AACATAGATAGATACATCAACACGACCTTCTCTGGTATTCTGTAAAGTATTGTGATGAC 780
 DB 721 AACATAGATAGATACATCAACACGACCTTCTCTGGTATTCTGTAAAGTATTGTGATGAC 780
 QY 781 AGTGCTTGCACCACTATGCAACTGATTTAGGAAAGACTGTTTATCAATCCACTGCTAAA 840
 DB 781 AGTGCTTGCACCACTATGCAACTGATTTAGGAAAGACTGTTTATCAATCCACTGCTAAA 840
 QY 841 TGTATTACAGATGTTCTCTGACAGATGCTACTTACTTGAACGACGAAACCCCTGGAC 900
 DB 841 TGTATTACAGATGTTCTCTGACAGATGCTACTTACTTGAACGACGAAACCCCTGGAC 900
 QY 901 AAGTCTTCTTTAGAGAGTCCATATTTCCTGTTAAGAAATCATGAACAATCCTTTTGA 960
 DB 901 AAGTCTTCTTTAGAGAGTCCATATTTCCTGTTAAGAAATCATGAACAATCCTTTTGA 960
 QY 961 GCCTATGACCACTATGAGGTGCTATGTTTATGCTTGTGTTGTTGTTGTTGTTGTTGTT 1020
 DB 961 GCCTATGACCACTATGAGGTGCTATGTTTATGCTTGTGTTGTTGTTGTTGTTGTTGTT 1020
 QY 1021 TCTTTTAGGCAATGTCAGAAATTTGATTGGCTCAGGGTTTATAGCTTTCTGTTGTTGTT 1080
 DB 1021 TCTTTTAGGCAATGTCAGAAATTTGATTGGCTCAGGGTTTATAGCTTTCTGTTGTTGTT 1080
 QY 1081 ATCTTCATGTTGCCCTGTTGCGGAACATTCATTACATGCTTAAAGCAACCCGCTGCTTC 1140
 DB 1081 ATCTTCATGTTGCCCTGTTGCGGAACATTCATTACATGCTTAAAGCAACCCGCTGCTTC 1140
 QY 1141 TCTTCATCTCCGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200
 DB 1141 TCTTCATCTCCGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200
 QY 1201 TCTTCATCTCTATAGAAATGCTGACGTTGGGAAACAGCTAAATAATTTATATAA 1254
 DB 1201 TCTTCATCTCTATAGAAATGCTGACGTTGGGAAACAGCTAAATAATTTATATAA 1254

RESULT 2
 ABS57951
 ID ABS57951 standard; DNA; 1512 BP.
 XX AC ABS57951;
 XX DT 04-FEB-2003 (first entry)
 XX DE S. pyogenes hyaluronate (HA) synthase gene, hasA.
 KW Gene; ds; hyaluronate synthase; HasA; HA; infection; phagocytic cell;
 KW polysaccharide capsule; hyaluronic acid; nonimmunogenic; macrophage;
 KW pharyngitis; impetigo; deep tissue infection; rheumatic fever;
 KW toxic shock syndrome; purity; cosmetic; eye surgery; viscosity;
 KW biocompatibility; vitreous fluid.
 XX OS Streptococcus pyogenes.
 XX FT Key Location/Qualifiers
 FT misc_signal 193..264

```

FT      /*tag= a
FT      /note= "Possible start of the HasA open reading utilising
FT      238...264
FT      misc_signal
FT      /*tag= b
FT      /note= "Possible start of the HasA open reading utilising
FT      250...264
FT      misc_signal
FT      /*tag= c
FT      /note= "Possible start of the HasA open reading utilising
FT      265...1452
FT      CDS
FT      /*tag= d
FT      /product= "Mature HasA protein"
FT      1488...1508
FT      CDS
FT      /*tag= e
FT      /product= "HasB protein"
FT      /partial
FT      /note= "No stop codon shown"
XX
XX      US6455304-B1.
XX
XX      24-SEP-2002.
XX
XX      03-SEP-1998; 98US-0146893.
XX
XX      01-JUL-1994; 94US-0270581.
XX
XX      (UYOK-) UNIV OKLAHOMA.
XX
XX      Weigel PH, DeAngelis PL, Papaconstantinou J;
XX
XX      WPI; 2003-066235/06.
XX      P-PSDB; ABG72066, ABG72067.
XX
XX      New purified nucleic acid segment encoding hyaluronate synthase, useful
XX      in recombinant DNA technology for the preparation of hyaluronic acid -
XX      Claim 2; Fig 7; 31pp; English.
XX
XX      The invention discloses a purified S. pyogenes nucleic acid encoding the
XX      hyaluronate (HA) synthase A (HasA). Streptococcal infections are a major
XX      health and economic problem and one reason for this is the ability of the
XX      bacteria to grow undetected by the body's phagocytic cells. They achieve
XX      this by coating themselves with polysaccharide capsules, such as
XX      hyaluronic acid capsules, which are generally nonimmunogenic and also
XX      prevent attachment of the bacteria to macrophages. Group A Streptococcus
XX      are responsible for numerous human diseases including pharyngitis,
XX      impetigo, deep tissue infections, rheumatic fever and toxic shock-like
XX      syndrome. The methods and compositions of the present invention are
XX      useful for recombinant DNA technology in the preparation of hyaluronic
XX      acid which has one or more improved properties, such as greater purity or
XX      ease of preparation, and produces larger amounts of relatively higher
XX      molecular weight and purity of hyaluronic acid than in prior art. High
XX      molecular weight HA has a wide variety of useful applications, ranging
XX      from cosmetics to eye surgery. In the case of the latter, its high
XX      viscosity and biocompatibility allows it to be a replacement for vitreous
XX      fluid. The sequence presented is the S. pyogenes HA synthase gene, hasA.
XX
XX      Sequence 1512 BP; 484 A; 227 C; 245 G; 556 T; 0 other;
SQ
Query Match      48.2%; Score 604.6; DB 25; Length 1512;
Best Local Similarity 68.1%; Pred. No. 2.6e-156;
Matches 846; Conservative 0; Mismatches 384; Indels 3; Gaps 1;
QY      10 TTAATAAACCTTCATACCTGTGGCCCTTTAGTATTTTGGGTACTGTTGATTTACGTC 69
DB      || ||||| ||| ||||| || ||||| || ||||| || ||||| || ||||| ||
DB      202 TTTAAAAAACTTTAAATGTTTATTTATCTCTTTTATTTTGAATCATCTCTGATTTATCTA 261
QY      70 AATGTTTATCTCTTTGGTGTCTAAAGGAAGCTGTGCAATTTATGCGCTTTTTCGCTGATAGCT 129
DB      ||| ||||| ||||| ||| ||| ||| ||||| ||| ||||| ||| ||||| |||
DB      262 AATATGATATCTATTGGAC---ATCAACTGTAGGATTTATGGAGTAATTAATTAAC 318

```

```

130 TACCTATTAGTCAAAATGCTCTTATCCCTTTTTCACAGCCATTTTAAAGGAAGGCTGGG 189
131 || || || || || || || || || || || || || || || || || || || || ||
319 TATCTAGTTATCAAACTTGGATATCTTCTTTATGAGCCATTTTAAAGGAATCCATC 378
190 CAATATAAGGTTGAGCCATTTATCCCTTTTATAAGGAAGATGCTGAGTCAATTTGCTAG 249
379 GACTATAAAGTTGCTGCTGTAATCTCTTCTTATTAAGGAATGCGAGTCAATTTAGAA 438
250 ACCTTAAAAGTGTTCAGACAAACCTTATCCCTTACGAGAAATTTATGTTGAGCAT 309
439 ACCTTAAAAGTGTGTTAGCAGACCTTATCCGTTATCAGAAATTTATATTTGTTGAT 498
310 GGAAGTGTGATGAGACAGGTATTAAGCGCATTTGAGAGCTATGTCGTGACATGCTGAC 369
499 GGGAGTTCAAACACAGATGCAATACAATTAATTCAGAGATGTAATATAGAGAGTGGAT 558
370 CTATCAAGCAATGCTATGTTTCATCGGTACAGAAAATCAAGGAAGCGTCTGACAG 429
559 ATTGTGGAAGCTTATGCTTACCGTTCCTTCTCAATAAAGGAACGCCATGCTCAA 618
430 GCCTGGGCTTTGAAAGATCAGACGCTGATGCTTTTTCACCGTTTGACTCAGATACCTAT 489
619 GCGTGGGCAITTTGAAGATCTGACGCTGACGTTTTTAAACGCTAGACTCAGATCTAT 678
490 ATCTACCTGTGCTTTTAGAGGAGTTGTTAAAAACCTTTTAAAGCCCACTGTTTGTCT 549
679 ATCTATCCAAATGCTTTAGAAGAACTCTCTAAAAGCTTCAATGATGACAGATTTATGCT 738
550 GCGACGGTGCACCTTATATGTCAGAAATAGACAAACCAACTCTTAAACCGTTGACAGAT 609
739 GCAACAGGACATTTGAATGCTAGAAACAGACAAACTAATCTATTAAACGCACTTACAGAT 798
610 ATTGCGTATGATAATGCTTTTGGCGTTGAAGAGCTGCCCAATCCGTTTACAGGTAATATC 669
799 ATCGGTTACGATTAATGCTTTGGGGTGGAGCGTCTCTCAATCAATTAACAGGTAATAT 858
670 CTGTTTGTCTAGGTCGCTTACGTTTACAGACGCGAGTGGTGTCTTACATAGAT 729
859 TTAGTTTGTCTAGGACCATTTAGTATTTATCGAGCTGAAGTATTTATCTTAACTTAGAG 918
730 AGATACATCAACAGACCTTCTCGGTGATTTCTCTAGTATTTGGTGATGACAGCTGCTTG 789
919 CGCTATAAAATCAAACTTCTAGTGTTCCTGTTAGCAATTGGGGATGATCGATGTTTA 978
790 ACCAACTATGCACTGATTTAGGAAGACTGTTTATCAATCCACTGCTAAATGATTAACA 849
979 ACAATATATGCTATGATTTAGGACGCACTGCTCAACCAATCAACAGCTAGATGTAATC 1038
850 GATGTTCTTGACAGATGCTCTTACTTGAAGCAGCAAAACCGCTGGAAACAGTCCCTTC 909
1039 GATGATACCTTTCCAAATTAAGAACTTATTTAAAGCAACAAATCGATGGAATAAATCTTT 1098
910 TTTAGAGAGTCCATTTATTTCTGTTAAGAAATCATGAACATCCCTTTGTAGCCCTATGG 969
1099 TTTAGAGATCTATTATTTCTGTTTAAATAATCTTTCTATCCCATCGTTGCTTATGG 1158
970 ACCATCTTGAAGGCTCTATGTTTATGATGCTGTTTATTCGTGTGGTGAATTTCTTTGA 1029
1159 ACTATTTTGAAGTCTGTTATGTTTATGATGTTGATGTTGTCGAATTTGGGAATCTTTGTT 1218
1030 GGCAATGTCAGAGAAATTTGATTTGGCTCAGGTTTTTACGCTTTCTGGTATTAATCTCAT 1089
1219 AATCAAGCTATTCAATATAGACCTTATTAACCTTTTTCGCTTTTATCCATCATCTTATC 1278
1090 GTTCGCTGTGCGAAGCAATTCATATGCTTAAGCAGCCCGCTGCTCTTCTTATCT 1149
1279 GTTCCTTTATGCTGTAATGTTCAATATATGTTGTAACATCTCTGCTAGTTTTTCTTATCT 1338
1150 CCGTTTATGGGCTGCTGCAATTTGTTTGTCTACAGCCCTTGAAATATATATCTCTTTTT 1209
1339 CCTCTGATGGAATATTACACTTGTGTTGCTTACAGCCCTTAAACTTATTTCTTTATGC 1398
1210 ACTATTAGAATGCTGACTGGGACACGCTAATAA 1244
QY

```



```

Db 1225 CCTCTGTATGGAATATACACTGTTGCTCTTACAGCCCTPAAACTTATCTTATGTC 1284
QY 1210 ACTATTAGAAATGCTGACTGGGGAACAGGTAAAAA 1244
Db 1285 ACCATTAAAAATACGGAATGGGGAACAGTAAAAA 1319

RESULT 4
ABL41013/c
ID ABL41013 standard; cDNA; 1653 BP.
XX AC
XX ABL41013;
XX 30-JUL-2002 (first entry)
XX DE Murine hyaluronan synthetase 2 (HAS2) cDNA fragment.
XX KW Hyaluronan synthetase 2; HAS2; tumour; gene therapy; mouse; gene; ss.
XX OS Mus sp.
XX PN JP2002065280-A.
XX PD 05-MAR-2002.
XX PF 01-SEP-2000; 2000JP-0266260.
XX PR 01-SEP-2000; 2000JP-0266260.
XX PA (SEK ) SEIKAGAKU KOGYO CO LTD.
XX DR WPI; 2002-397661/43.
XX PT A DNA vector for inhibiting the formation of a malignant tumor -
XX PS Claim 1; Page 6; 9pp; Japanese.
XX CC The invention relates to a DNA vector containing a 20 nucleotide sequence
CC of the 3'-end of a mouse hyaluronan synthetase 2 (HAS2) cDNA sequence,
CC or DNA hybridisable with a polynucleotide complementary to it under
CC stringent conditions, and which can inhibit the translation of a
CC hyaluronan synthetase 2 gene in which the transcription initiating point
CC is recombined 5' upstream of the DNA. The vector is used for inhibiting
CC the formation of a malignant tumour. The present sequence represents a
CC mouse HAS2 cDNA fragment.
XX SQ Sequence 1653 BP; 471 A; 387 C; 353 G; 442 T; 0 other;

Query Match 4.9%; Score 61.6; DB 24; Length 1653;
Best Local Similarity 49.1%; Pred. No. 1.4e-06;
Matches 22; Conservative 0; Mismatches 224; Indels 6; Gaps 2;

QY 474 TGACTCAGATACCTATATACCTACCCCTGCTTTAGAGAGGTGTTAAACACCTT---TAA 530
Db 1024 TGACTCAGATATGCTTGACCCCTCCTCATCTGTGGAGATGTTGAGGCTCTAGAGGA 965
QY 531 TGACCAACACTGTTTTGCTGGACGGGTCACTTAATGTCAGAAATAGACAAACCAATCT 590
Db 964 AGACCTATGTTGGAGAGATGTCAGATTTTAAACAGTATGATTCG 905
QY 591 CTTAACACGCTTGACAGATATGCTATGCTATGCTTTTGGCGTTGACAGCGTCCCA 650
Db 904 GATCTCTCTTCCTCAGCAGCGTGAGATCTGGATGCTTTTAAATAGAGAGGCGCTGCCA 845
QY 651 ATCGGTTACAGTAATATCTCTGTTGCTCAGGTCGCTACGTTTACAGACGCGAGGT 710
Db 844 GTCTTATTTTGGCTGTCTCCAGTGTCATAGCGGCTCTCGGGAATGTACAGAAACCTCT 785
QY 711 GGTGTGTTCTACATAGATAGATACATCAACGACCTTCCCTGGGTATTCCTGTAGTAT 770
Db 784 GCTGCATGAATTTGTGGAAGACTGGTACATCAGGAATTCATGGGTAACCAACGAGTTT 725

```

```

QY 771 TGGTGATGACAGTGCTTGACCACTATGCAACTGATTAGGAAAGACTGTTTATCAATC 830
Db 724 TGGTGACGACAGGACCTTACCAACAGGGTGTGAGTCTGGGCTATGCAACTAAATACAC 665
QY 831 CACTG---CTAAATGTATTACAGATGTTCTGCACAAGATGCTTACTTACTTGAAGCAGCA 887
Db 664 GGCTCGGTCCAAGTGCCTTACTGAACTCCCATAGATATCTGAGATGGCTGAACACGCA 605
QY 888 AAACCGCTGGAAAGTCTTCTTTTAGAGAGT 919
Db 604 GACCCGTTGGAGCAAGTCTTACTTCGAGAGT 573

RESULT 5
ABZ76735
ID ABZ76735 standard; cDNA; 1659 BP.
XX AC
XX ABZ76735;
XX 01-APR-2000 (first entry)
XX DE Mouse hyaluronan synthase 2 encoding cDNA SEQ ID NO:2.
XX KW Mouse; hyaluronan synthase; HAS; HAS2; enzyme; ophthalmological;
XX KW osteopathic; antiarthritic; gene therapy; angiogenesis inhibitor;
XX KW hyaluronic acid; dihydrazide; bioconjugate; dry eye syndrome; eye;
XX KW osteoarthritis; gene; ss.
XX OS Mus sp.
XX FH Key
XX FT Location/Qualifiers
XX FT 1..1659
XX FT /tag= a
XX FT /product= "hyaluronan synthase 2 (HAS2)"
XX PN WO2003006068-A1.
XX PD 23-JAN-2003.
XX PF 10-JUL-2001; 2001WO-US21785.
XX PR 10-JUL-2001; 2001WO-US21785.
XX PA (CLEA-) CLEAR SOLUTIONS BIOTECH INC.
XX PI Dehazya P, Chen W;
XX DR WPI; 2003-221664/21.
XX DR P-PSDB; ABP96029.
XX PT Novel dihydrazide derivatized hyaluronic acid/nucleic acid bioconjugate
XX PT for treating dry eye syndrome, has derivatized hyaluronic acid
XX PT crosslinked to nucleic acid encoding protein with hyaluronan synthase
XX PT activity -
XX PS Claim 19; Page 56-57; 62pp; English.
XX CC The present invention describes a dihydrazide derivatised hyaluronic acid
XX CC (HA)/nucleic acid (NA) bioconjugate (I), with derivatised HA linked to NA
XX CC encoding protein with hyaluronan synthase (HAS) activity, where NA has
XX CC sequence with 79-85% sequence identity (SI) to a sequence of 1752, 1659
XX CC or 1665 nucleotides (see ABZ76734 to ABZ76736) or a sequence encoding
XX CC proteins with 66-90% SI to a sequence of 583, 552 or 554 amino acids (see
XX CC ABP96028 to ABP96030). (I) has ophthalmological, osteopathic and
XX CC antiarthritic activities, and can be used in gene therapy and as an
XX CC inhibitor of angiogenesis, and as an inducer of expression of (HA) in
XX CC human corneal epithelial cell. (I) is useful for transfecting a cell of
XX CC an eye with nucleic acid, where the nucleic acid comprises (ABZ76734 to
XX CC ABZ76736), by contacting the cell with (I) comprising the nucleic acid.
XX CC (I) is useful for treating dry eye syndrome in an individual. (I) is
XX CC useful in gene therapy applications for the treatment of a variety of
XX CC medical conditions including dry eye syndrome or other medical conditions
XX CC where an increase in the production of (HA) in the eye would be

```


52 15919ACGACAGGCACTTACCAACAGGGTGTGTGAGTCCTGGGCTATGGCAACTAATATACAC 1493


```

Db      1220 AGACCCATGTTGGAGGTGTTGGAGGAGATGTCAGATTTTAAACAAGTAATTCCTG 1279
QY      591 CTTACACGGCTGACAGATATCCCTATGATTAATGCTTTTGGCGTTGAACGAGCTGCCCA 650
Db      1280 GATCTCCTCTCAGCAGCGTGAGATAGTGGATGCTTTTAAATATGAAGGCGCTGCCA 1339
QY      651 ATCCGTTACAGGTAATATCCCTGTTTCTCAGTCGCTTAGCGTTTACAGACGCGAGGT 710
Db      1340 GTCATATTTGGCTGTGTCAGTGTCATTAAGCGTCTCTGGGAATGTACAGAACTCGTT 1399
QY      711 GGTGTTCTTACATAGATAGATACATCAACAGACCTTCTCTGGTATTCCTGTAAGTAT 770
Db      1400 GTCATCAATTTTGTGAAGACTGGTACAACTCAGGAATTCATGGTAACCAATGCACTT 1459
QY      771 TGGTGATGACAGGTGTTGACCAACTATGCACTGATTTAGGAAGACGTTTATCAATC 830
Db      1460 TGGTGACGACAGCCCTTACCAACAGGGTGTGAGTCTGGGCTATGCAACTAAATACAC 1519
QY      831 CACTG---CPAATGTAATACAGATGTTCTCAGCAAGATGTCATCTACTTACTTGAAGCAGCA 887
Db      1520 GGCTCGGTCCAAGTGCCTTACTGAACTCCCATAGATATCTGAGATGCTGACACGCA 1579
QY      888 AAACCGTGAACAGTCTCTTTAGAGAGT 919
Db      1580 GACCGTGGAGCAAGCTCTACTCCGAGAGT 1611

```

RESULT 9

AAA14914

ID AAA14914 standard; DNA; 7193 bp.

XX AC AAA14914;

XX DT 08-AUG-2000 (first entry)

XX DE DNA encoding nodulation efficiency factors.

XX KW Nodulation efficiency factor; Sinorhizobium meliloti USDA 1170; nodJ;
 KW strain NRG 185; cut-leaf medic; nod gene; nodA; nodB; nodC; nodD; nodI;
 KW nodulation factor; nodulation efficiency; Medicago lacinata;
 KW nitrogen fixation; legume; ss.

XX OS Sinorhizobium meliloti.

XX FH Key Location/Qualifiers
 FT CDS complement (3..275)

FT FT /*tag= a

FT FT /product= "nodD"

FT FT /note= "encodes AAY84842"

FT FT 508..511

FT FT /*tag= b

FT FT 526..1116

FT FT /*tag= c

FT FT /product= "nodA"

FT FT /note= "encodes AAY84843"

FT FT 1100..1103

FT FT /*tag= d

FT FT 1113..11766

FT FT /*tag= e

FT FT /product= "nodB"

FT FT /note= "encodes AAY84844"

FT FT 1766..1769

FT FT /*tag= f

FT FT 1781..3046

FT FT /product= "nodC"

FT FT /*tag= g

FT FT /trans_except= (pos: 1961..1963, aa: Glu)

FT FT /note= "encodes AAY84845"

FT FT 3070..3074

FT FT /*tag= h

FT FT 3092..4138

FT FT /*tag= i

FT FT /product= "nodI"
 FT FT /note= "encodes AAY84846"
 FT FT 4125..4127
 FT FT /*tag= j
 FT FT 4142..4930
 FT FT /*tag= k
 FT FT /product= "nodJ"
 FT FT /note= "encodes AAY84847"
 FT FT 5075..5079
 FT FT /*tag= l
 FT FT 5103..5321
 FT FT /*tag= m
 FT FT /product= "ORF1 protein"
 FT FT /note= "encodes AAY84848"
 FT FT 5353..5358
 FT FT /*tag= n
 FT FT 5368..5706
 FT FT /*tag= o
 FT FT /product= "ORF2 protein"
 FT FT /trans_except= (pos: 5533..5535, aa: Gln)
 FT FT /note= "encodes AAY84849"
 FT FT complement (5866..6297)
 FT FT /*tag= p
 FT FT /product= "ORF3 protein"
 FT FT /note= "encodes AAY84850"
 FT FT 6294..6297
 FT FT /*tag= q
 FT FT complement (6306..6563)
 FT FT /*tag= r
 FT FT /product= "ORF4 protein"
 FT FT /note= "encodes AAY84851"
 FT FT 6559..6564
 FT FT /*tag= s
 FT FT 6567..6572
 FT FT /*tag= t
 FT FT complement (6613..7118)
 FT FT /*tag= u
 FT FT /product= "ORF5 protein"
 FT FT /note= "encodes AAY84852"
 FT FT 7113..7117
 FT FT /*tag= v

XX WO200022138-A1.

XX PD 20-APR-2000.

XX PF 13-OCT-1999; 99WO-CA00955.

XX PR 14-OCT-1998; 98US-0104162.

XX PA (AGRI-) AGRIC & AGRI-FOOD CANADA.

XX PI Barran LR, Bromfield ESP, Brown DCW;

XX DR WPI; 2000-317991/27.

XX DR P-PSDB; AAY84842, AAY84843, AAY84844, AAY84845, AAY84846, AAY84847,

XX DR AAY84848, AAY84849, AAY84850, AAY84851, AAY84852.

XX PT Isolated nodulation efficiency factor for improving nitrogen fixation
 in legumes comprising a portion of EcoRI/BamHI fragment of
 Sinorhizobium meliloti that confers the nodulating ability -

XX PS Claim 1; Fig 1; 47pp; English.

XX CC The present sequence encodes nodulation efficiency factors. The present
 sequence is a 7.2 kb EcoRI/BamHI fragment from Sinorhizobium meliloti
 USDA 1170, that gives S. meliloti strain NRG 185 the ability to nodulate
 at least 50% of inoculated Medicago lacinata (cut-leaf medic) plants
 within 10 days of inoculation. The fragment contains nod genes nodA,
 nodB, nodC, nodD, nodI, nodJ. The nodulation factors are used to
 increase the nodulation efficiency of Sinorhizobium for Medicago
 lacinata. The nodulation factors are also used to improve nitrogen
 fixation in legumes.

XX	SQ	Sequence	7193 BP; 1434 A; 2094 C; 2132 G; 1532 T; 1 other;
Query Match		4.5%;	Score 56; DB 21; Length 7193;
Best Local Similarity		46.6%;	Pred. No. 9.1e-05;
Matches 179; Conservative		0; Mismatches 205; Indels 0; Gaps 0;	
QY	412	GGAAAGCGTCATCGACAGCGCTGGCGCTTTGAAGATCAGACGCTGATGCTTTTGGACC	471
DB	2138	GGAAAGCGGAAGCGCAGATTGGCGCGATAGGTCAATCCTCTGGGAACCTGGTGCTGAAT	2197
QY	472	GTTGACTCAGATCTATATCTACCTCGATGCTTTAGAGGAGTTCGTTAAAAACCTTTAAT	531
DB	2198	GTGACTCAGACAGACGATTCGATGTCGTCTCCAGCTTTGCTCGAAGATGCGA	2257
QY	532	GACCAACTGTTTTGTCGCGACGGGTCACTTAATGTCAGAAATAGACAAACCAATCTC	591
DB	2258	AATCCAGAGGTTGGTGGCGCATGGGTCACTCAGCGCTAGCAATCGGAGTGACACCTGG	2317
QY	592	TTACACGCTTGACAGATATTCGCTATGATATATGCTTTTGGCGTTGACGAGCTGCCAA	651
DB	2318	CTGACGAGATTGATCGATGAGTATTTGGCTTGCCTGCAACGAAGACGCGCGCACAG	2377
QY	652	TCCGTTACAGGTAATATCCTTGTGTTGCTCAGGCTAGCGTTAGCGTTTACACGCGAGGTG	711
DB	2378	GTCGCTTCGGTGGCGGTATGTTGTCTCGGCGCCATGTGCTATGATCCGTCGGTCGCG	2437
QY	712	GTTGTTCTTAACATAGATAGATACATCAACACGACCTTCCTCGGTATTCCTGTAAGTAT	771
DB	2438	CTCGCTGGCTGTTGACCATGACGAGAAACGACGCTGTTTCGGGTAAACCAACGACCTC	2497
QY	772	GGTGATGACAGGTGCTTGACCAAC	795
DB	2498	GGTGAGGATGCTCATCTGACGATC	2521
RESULT 10			
AAV30458/c			
ID	AAV30458	standard; DNA; 534720 BP.	
XX	AAV30458;		
XX	14-OCT-1998	(first entry)	
XX	Rhizobium species plasmid pNGR234a.		
XX	Symbiosis; open reading frame; ORF; plasmid; vector; transportation;		
KW	degradation; metabolism; host range; nitrogen fixation; modulation;		
KW	legume; plant; ds.		
XX	Rhizobium sp.		
XX	Key	Location/Qualifiers	
FF	CDS	417796..418671	
FT		/*tag= a	
FT		/*standard_name= "ORF K1"	
FT		/*product= "oligopeptide permease"	
FT		/*note= "homologous to the oppc gene"	
FT	CDS	418673..419680	
FT		/*tag= b	
FT		/*standard_name= "ORF K2"	
FT		/*product= "oligopeptide permease"	
FT		/*note= "homologous to the oppd gene"	
FT	CDS	419677..420738	
FT		/*tag= c	
FT		/*standard_name= "ORF K3"	
FT		/*product= "oligopeptide permease"	
FT		/*note= "homologous to the oppf gene"	
FT	CDS	420774..422159	
FT		/*tag= d	
FT		/*standard_name= "ORF K4"	
FT		/*product= "encapsulation-like protein"	
FT		/*note= "homologous to the capA gene"	


```

PF 10-JUL-1997; 97WO-IB00950.
XX
XX 20-MAY-1997; 97GB-0010395.
PR 12-JUL-1996; 96EP-0730001.
XX
XX (MOLE-) INST MOLECULAR BIOTECHNOLOGY.
PA (BIOL-) LAB BIOLOGIE MOLECULAIRE PLANTES SUPERIE.
XX
PI Broughton WJ, Freiberg CB, Perret XP, Rosenthal A;
XX
XX WPI; 1998-110606/10.
DR
XX
XX New isolated symbiotic plasmid from Rhizobium sp. NGR234 - used to
PT develop products for modifying plant characteristics, e.g. nitrogen
PT fixation, synthesis of compounds and stress response
XX
XX Claim 1; Fig 3; 228pp; English.
XX
XX This is the nucleotide sequence of the plasmid pNGR234a isolated from
CC Rhizobium sp. NGR234. Open reading frames (ORF) derivable from the
CC nucleotide sequence are claimed. The nucleotide sequences or ORFs can
CC be used e.g. in the transportation of compounds to and from an organism
CC which is a host to at least one of the nucleotide sequences, ORFs or
CC proteins, the degradation and/or metabolism of organic, inorganic,
CC natural or xenobiotic substances in a host organism or the modification
CC of the host range, nitrogen fixation abilities; for obtaining a synthetic
CC minimal set of ORFs required for functional Rhizobium-legume symbiosis,
CC especially for nodulation efficiency on host plants.
XX
XX Sequence 536165 BP; 111291 A; 155755 C; 157864 G; 111255 T; 0 other;
XX SQ
Query Match 4.3%; Score 53.8; DB 19; Length 536165;
Best Local Similarity 46.1%; Pred. No. 0.0022;
Matches 181; Conservative 0; Mismatches 212; Indels 0; Gaps 0;
QY 403 AAAATPCAAAGAGCGTCATGCACAGCGCTGGGCGCTTTGAAAGATCAGACGCTGATGTC 462
DB 156991 AAGAAGCGTCGGAAGCGCAAGCGCAGATTGCGGATACGGGAATCATCGGAGATTG 156932
QY 463 TTTTGACCGTGTACTCAGTACTTATCTATCTACCGTGATGCTTAGAGGAGTGTAAAA 522
DB 156931 GTGCTCAACGTGTACTCGGACACAGCACCATTGCGCGGAGGTAGTCACGAAACTGCCCTG 156872
QY 523 ACCTTAAATGACCAACTGTTTGTCTGCGACGGGTGTCACCTTAATGTCAAGAAATAGACAA 582
DB 156871 AAGATGTACAGTCCGCGGTGCGCGCGATGGGTGCTGACGCGCCACACCGCAGC 156812
QY 583 ACCAATCTTTAACAGCGTTGACAGATATTCGCTATGATATGCTTTTGGCGTTGAACA 642
DB 156811 GACATAGCGTGAAGCGGTGTATCGACATGAGTACTGGCTGCGCTGCAACGAGGAACGC 156752
QY 643 GCTGCCCAATCCGTTACAGGTAAATCTTGTCTGACGTCGCTTAGCGTTTACAGA 702
DB 156751 GCAGCACAGGTCGCTTGGACCGCTATGCTGTGCGCGCGGTGTCGCAATGATACGG 156692
QY 703 CGGAGAGTGTGTTGCTTCAATATAGATAGATACATCAACAGACCTTCCTGGGTAFTCCT 762
DB 156691 CGGTCGCACTTCCTATTGCTGCTGATAAATACGAGACCACTGTTTTCAGGACGGCA 156632
QY 763 GTAAGTATTGGTGATGACAGTGCTTGACCAAC 795
DB 156631 ACAGACTTCGGGGAGACCGCACTCACAATC 156599
RESULT 12
AAV18822
ID AAV18822 standard; cDNA; 1665 BP.
XX
XX AAV18822;
AC
XX
XX 11-JUN-1998 (first entry)
DT
XX
XX Murine hyaluronan synthase-3 cDNA.

```

```

XX
XX Murine; hyaluronan synthase-3; HS-3; hyaluronan;
XX hyaluronic acid; ss.
XX
XX Mus sp.
XX
XX Key Location/Qualifiers
XX CDS 1..1665
XX FT /*tag= a
XX FT /product= hyaluronan_synthase-3
XX
XX WO9800551-A2.
XX
XX 08-JAN-1998.
XX
XX 03-JUL-1997; 97WO-US11761.
XX
XX 05-MAR-1997; 97US-0812008.
XX 03-JUL-1996; 96US-0675499.
XX
XX (MAYO-) MAYO FOUNDATION.
XX
XX Augustine ML, McDonald JA, Spicer AP;
XX
XX WPI; 1998-086976/08.
XX P-PSDB; AAW50010.
XX
XX Nucleic acid encoding hyaluronan synthase-2 or -3 or their active
XX fragments - used for treating arthritis, in wound repair, for
XX delivering drugs etc.
XX
XX Claim 11; Pages 75-78; 115pp; English.
XX
XX The present sequence encodes murine hyaluronan synthase-3 (HS-3),
XX which can be used to alter hyaluronan, i.e. hyaluronic acid (HA),
XX production, or for in vitro synthesis of HA.
XX
XX HA can be used to treat arthritis or perforated ear drums, protect
XX eyes during surgery, deliver drugs, prevent post-operative scarring
XX or adhesion formation, heal wounds and prevent restenosis following
XX balloon angioplasty. Modulation of HA production in vivo may be
XX useful in, e.g. Graves disease, mesothelioma, Wilms' tumour and
XX oedema associated with inflammation of lung and kidney, all of
XX which are associated with elevated levels of HA in tissues or
XX serum.
XX
XX Sequence 1665 BP; 340 A; 453 C; 421 G; 421 T; 0 other;
XX SQ
Query Match 3.9%; Score 49.4; DB 19; Length 1665;
Best Local Similarity 49.2%; Pred. No. 0.0033;
Matches 159; Conservative 0; Mismatches 161; Indels 3; Gaps 1;
QY 600 CTTGACAGATATTCGCTATGATATGCTTTTGGGCTTGAACGAGCTGCCAATCGGTAC 659
DB 774 CCTGAGCAGTGTGAGGTACTGGATGGCTTTCAACGTGGAGCGGGCTGCCAGTCTACTT 833
QY 660 AGGTAAATATCTTGTGTTGCTCAGGTCCGCTTACAGCTTTTACAGACGAGGTGTTGTTCC 719
DB 834 TGGCTGTGCAATCATATGTTAGTGGGCTTTGGGCATGTACCGCACAGCCTCTCAGCA 893
QY 720 TAACATAGATAGATACATCAACGACACCTTCCTGGGTATTTCTGTGAAGTATTTGGTGATGA 779
DB 894 GTTCTCGGAGGATTTGGTACCATCAGAAGTCTCTAGGCAGCAAGTGCAGCTTTGGGGATCA 953
QY 780 CAGTGCTTTGACCACTATGCACTGATTTAGGAAA---GACTGTTTATCAATCCACTGC 836
DB 954 TCGCACCTTACCAACCGGAGTCTCTGAGTCTTGGCTACCGGACTTAAGTATACAGCAGCTC 1013
QY 837 TAAATGTATTACAGATGTTCTGACAAGATGTCCTACTTACTTGAGCAGCAAAACGCTG 896
DB 1014 TAAGTGCTTCACAGAGACCCCTAGTACTTCGATGGCTCAATACAGCAAAACGCTG 1073
QY 897 GAACAAGTCTCTTCTTTAGAGAGT 919

```

[illegible]

CC degeneration or genes related to lipid biosynthesis that helps to restore
 CC the lipid component of the tear film, and as reagents for in vitro
 CC transformation of any cell, preferably a eukaryotic cell, more preferably
 CC a human eye cell. The present sequence encodes mouse HAS3 which is used
 CC in the exemplification of the present invention.

XX
 SQ Sequence 1665 BP; 340 A; 453 C; 451 G; 421 T; 0 other;

Query Match 3.9%; Score 49.4; DB 25; Length 1665;
 Best Local Similarity 49.2%; Pred. No. 0.0033;
 Matches 159; Conservative 0; Mismatches 161; Indels 3; Gaps 1;
 QY 600 CTTGACAGATATCGCTATGATATGCTTTGGCGTTGACGAGCTGCCAATCCGTTAC 659
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 774 CTGAGCAGCTGAGGTACTGATGGCTTCAACGGGAGCGGCCCTGCCAGTCCCTACTT 833
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 660 AGTAAATATCCTTTGCTCAGGTCGCTAGCTTTACAGCGCGAGGTGTTGTTCC 719
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 834 TGGCTGTGCAATGATTAGTGGCCCTTTGGCAGTGATCCGCAACAGCCTCCTTCAGCA 893
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 720 TAACATAGATAGATACATCAACGAGCCTTCCTGGGTATTCCTGTAGTATTGGTGATGA 779
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 894 GTTCCTGGAGGATTCGATCACCATCAGAAGTTCCTAGCAGCAAGTGCAGCTTTGGGGATGA 953
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 780 CAGGTGCTTGACCAACTATGCAACTGATTAGGAAA--GACTGTTTATCAATCCACTGC 836
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 954 TCGGCACCTTACCAACGAGTCTCTGAGTCTTGCTACCGGACTAAGTATACAGCAGGCTC 1013
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 837 TAAATGTTATACAGATGTTCCGCAAGATGCTTACTTACTTGAAGCAGCAAAACCGCTG 896
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1014 TAAATGCTCAGAGAGACCCCACTAGTACTTCGATGGCTCAATCAGCAAAACCGCTG 1073
 QY 897 GACAACTCCTTCTTAGAGAGT 919
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1074 GACAACTTACTTTCGGGAAT 1096
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15

AA10595/C
 ID AA10595 standard; DNA; 6741 BP.
 XX
 AC AA10595;
 XX
 DT 29-JUN-2000 (first entry)
 XX
 DE Gene encoding a subunit of cellulose synthase.
 XX
 KW Cellulose synthase; cellulose production; increase yield; ds.
 XX
 OS Vigna angularis.
 XX
 PN JP2000060568-A.
 XX
 PD 29-FEB-2000.
 XX
 PF 26-AUG-1998; 98JP-0239998.
 XX
 PR 26-AUG-1998; 98JP-0239998.
 XX
 PA (MIZUO) MIZUNO K.
 PA (OJIP) OJI PAPER CO.
 XX
 DR WPI; 2000-342371/30.
 DR P-PSDB; AAY85180.
 XX

PT A gene encoding a cellulose synthetic equipment - for the improvement
 PT in the amount of cellulose synthesised in a plant body
 XX
 PS Claim 2; Page 22-31; 32pp; Japanese.
 XX
 CC This sequence represents a gene encoding a subunit of the cellulose
 CC synthase complex of Vigna angularis. The invention relates to subunits of
 CC cellulose synthetic equipment, that can be used to increase the amount of

CC cellulose synthesised by a plant. The proteins and genes encoding them
 CC can also be used to improve the properties of the cellulose being
 CC produced by a plant.

XX
 SQ Sequence 6741 BP; 1712 A; 870 C; 1468 G; 1374 T; 1317 other;

Query Match 3.8%; Score 47.8; DB 21; Length 6741;
 Best Local Similarity 18.4%; Pred. No. 0.016;
 Matches 114; Conservative 201; Mismatches 301; Indels 2; Gaps 1;
 QY 600 CTTGACAGATATTCCTATGATATGCTTTGGGTTGACGAGCTGCCAATCCGTTAC 659
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 6167 STNSTBYRANSTTFCNCYRASSTTTTCSRAKYNCCDCYNSRNCSTCCYCY 6108
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 660 AGTAATATCCCTTTGCTCAGGTCGCTAGCGTTTACAGACGCGAGGTGTTGTTCC 719
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 6107 SSRSRNCCTTSSRCTBYRAYACSRSTSTDSRRSRNCSTTCTBTCTTCTTTTSTTTNST 6048
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 720 TAACATAGATAGATACATCAACGAGCCTTCTGGGTATTCCTGTAGTATTGGTGATGA 779
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 6047 CDNSTBDRCDRSGRSYSSRRCSTCTCTYAKTBTBTYSYDAYDARDRCRDYDASRC 5988
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 780 CAGTGTCTTGACCAACTATGCAACTGATTAGGAAAGACTGTTTATCAATCCACTGTA 839
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 5967 STSYSDAYSTTTNSTSRBTBNTSTNSDTSRTTSRTTYYDACCYTCNSRCYTDYR 5928
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 840 ATGATTTACAGATGTTCTGACAGATGTCATCTTACTTGAACGAGCAAAACCGCTGAA 899
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 5927 CTBYRATBNCSTNSTAKRNCSTSRCTTSRNCCTYSCCNSTDNDCTBDRCYRCSRY 5868
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 900 CAAGTCTCTTTAGAGAGTCCATTTATCTGTTAAGAAATCAT--GNAACAATCCTTTT 957
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 5867 SNCRYSISTYDATTBTSRSGISNSTRCDYDAYSDYDASTSDTTSRSDTTTCYFTSN 5808
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 958 GTAGCCCTATGACCACTATGAGGTGCTCTATGTTTATGATGCTTGTATTTCTGTGTTG 1017
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 5807 STSRDNRCTTCYSSRCYTSYCTSYDYRAAKCYTYSRNSNSTSYDADNSTTB 5748
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1018 GATTTCTTTAGGCAATGTCAGAGATTTGATGCGTCAAGGCTTTTACCTTTCTGCTG 1077
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 5747 STTTCYTTTTSTNCYNSYRANSTTBYRACYRACYSCYCTYSYDASTCYTSYSG 5688
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1078 ATTATCTTCATTTGCTGCTGCGAATCATGCTTACATGCTTAAGCACCCTGCTCC 1137
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 5687 TBCYTTBYSTBNCRCNCYSSTYRANCYSSTYDABNSTYSTSTCTBYRAYSYDAYD 5628
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1138 TCTCTGTATCTCCGTTTATGGGGTCTGCAATTTGTTGCTCTACAGCCCTGAAATTA 1197
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 5627 ATBDCNSTRCTCTTTTTCNSTRYCAKTNSTNCTYTRCYSRCSYSYSTBT 5568
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1198 TATTCCTCTTTTACTATT 1215
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 5567 TTTTCTTTTCTTTTCTTTT 5550
 Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Search completed: October 2, 2003, 09:10:06
 Job time : 385 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2003, 09:01:58 : Search time 81 Seconds
(without alignments)
6833.265 Million cell updates/sec

Title: US-09-469-200D-1
Perfect score: 1254
Sequence: 1 atggaacattataaaacct.....cagctataaaattattataa 1254

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	604.6	48.2	1512	2	US-08-270-581-1
2	604.6	48.2	1512	4	US-09-146-893-1
3	61.6	4.9	2947	4	US-08-675-499A-1
4	53.8	4.3	536165	4	US-09-214-808-1
5	45.6	3.6	7218	1	US-08-232-463-14
6	37.8	3.0	2173	1	US-08-036-210-6
7	37.8	3.0	2173	2	US-08-449-609-6
8	37.8	3.0	2173	4	US-09-361-096A-6
9	37.8	3.0	2309	1	US-08-036-210-10
10	37.8	3.0	2309	2	US-08-449-609-10
11	37.8	3.0	2309	4	US-09-361-096A-10
12	37.8	3.0	2692	1	US-08-036-210-14
13	37.8	3.0	2692	2	US-08-449-609-14
14	37.8	3.0	2692	4	US-09-361-096A-14
15	37.8	3.0	3973	1	US-08-036-210-21
16	37.8	3.0	3973	2	US-08-449-609-21
17	37.8	3.0	3973	4	US-09-361-096A-21
18	37	3.0	1257	4	US-09-134-001C-1214
19	35.6	2.8	19124	4	US-08-487-826B-13
20	35.4	2.8	580073	4	US-08-545-528D-1
21	35	2.8	1371	4	US-09-328-352-3464
22	35	2.8	1664976	4	US-08-916-421B-1
23	34.6	2.8	1392	4	US-09-328-352-474
24	34.4	2.7	1830121	4	US-09-557-884-1
25	34.4	2.7	1830121	4	US-09-643-990A-1
26	34.2	2.7	1866	1	US-08-328-961-1
27	34.2	2.7	1866	1	US-08-462-397-1

c	28	34.2	2.7	2310	1	US-08-417-330A-15	Sequence 15, Appl
	29	34.2	2.7	2848	2	US-08-805-918-1	Sequence 1, Appl
	30	34.2	2.7	4106	2	US-08-702-572-14	Sequence 14, Appl
	31	34.2	2.7	4732	6	5521093-4	Patent No. 5521093
c	32	34.2	2.7	1664976	4	US-08-916-421B-1	Sequence 1, Appl
c	33	34.2	2.7	2169	4	US-09-434-408-3	Sequence 3, Appl
	34	33.6	2.7	2945	2	US-08-687-372-2	Sequence 2, Appl
	35	33.6	2.7	9060	3	US-08-378-313-20	Sequence 20, Appl
	36	33.4	2.7	2184	4	US-09-134-001C-1230	Sequence 1230, Appl
c	37	33.4	2.7	6755	3	US-08-931-999-4	Sequence 4, Appl
	38	33.2	2.6	2628	2	US-08-696-944-1	Sequence 1, Appl
c	39	33	2.6	2164	3	US-08-760-615-3	Sequence 1, Appl
c	40	33	2.6	11460	4	US-09-336-910A-1	Sequence 1, Appl
	41	32.8	2.8	65042	4	US-09-784-316-3	Sequence 3, Appl
	42	32.4	2.6	799	3	US-08-961-083-13	Sequence 13, Appl
	43	32.4	2.6	799	4	US-09-536-784-13	Sequence 13, Appl
	44	32.4	2.6	1242	4	US-09-328-352-101	Sequence 101, Appl
	45	32.4	2.6	1404	4	US-09-134-001C-2464	Sequence 2464, Appl

ALIGNMENTS

RESULT 1
US-08-270-581-1
: Sequence 1, Application US/08270581
: Patent No. 5856168
: GENERAL INFORMATION:
: APPLICANT: Weigel, Paul H.
: APPLICANT: DeAngelis, Paul L.
: APPLICANT: Papaconstantinou, John
: TITLE OF INVENTION: Hyaluronate Synthase Gene and Uses
: TITLE OF INVENTION: Thereof
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Arnold, White & Durkee
: STREET: P. O. Box 4433
: CITY: Houston
: STATE: TX
: COUNTRY: USA
: ZIP: 77210-4433

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/270,581
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UTSG:161\PAR
TELEPHONE: 713-787-1400
TELEFAX: 713-789-2679
INFORMATION FOR SEQ ID NO: 1:
LENGTH: 1512 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 193..1449
FEATURE:
NAME/KEY: CDS
LOCATION: 1488..1510
US-08-270-581-1

Query Match 48.2% Score 604.6; DB 2; Length 1512;

1219 AATCAAGCTATTCAATTAGACCTTATTAAACCTTTTGGCTTTTATTCATCATCTTTATC 1278
 QY 1090 GTTGGCCCTGTGTGGGAACATTCATTACATGCTTAAGCACCCGCTGCTTCTTGTTACT 1149
 Db 1279 GTTGGCTTATGTCGTAATGTTTCATATATGGTCAACATCCTGCTAGTTTTGTATCT 1338
 QY 1150 CGGTTTATGGGTGCTGCATTTGTTTGCTTCACAGCCCTGAAATATATATCTCTTTT 1209
 Db 1339 CCTCTGATAGGAATATACACTTGTGTCTTACAGCCCTAAACCTTATCTTTATGC 1398
 QY 1210 ACTATTAGAATGCTCACTGGGGAACACGTA AAAA 1244
 Db 1399 ACATTAAATACGAATGGGAACACGTA AAAA 1433

 RESULT 2
 US-09-146-893-1
 ; Sequence 1, Application US/09146893
 ; Patent No. 6455304
 ; GENERAL INFORMATION:
 ; APPLICANT: Weigel, Paul H.
 ; deAngelis, Paul L.
 ; Papaconstantinou, John
 ; TITLE OF INVENTION: Hyaluronate Synthase Gene and Uses
 ; Thereof
 ; NUMBER OF SEQUENCES: 7
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Arnold, White & Durkee
 ; STREET: P. O. Box 4433
 ; CITY: Houston
 ; STATE: TX
 ; COUNTRY: USA
 ; ZIP: 77210-4433
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/146,893
 ; FILING DATE: 03-Sep-1998
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/270,581
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Parker, David L.
 ; REGISTRATION NUMBER: 32,165
 ; REFERENCE/DOCKET NUMBER: UTSG:161\PAR
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 713-787-1400
 ; TELEFAX: 713-789-2679
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1512 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 193..1449
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1488..1510
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-146-893-1

 Query Match 48.2%; Score 604.6; DB 4; Length 1512;
 Best Local Similarity 68.7%; Pred. No. 7.4e-169;
 Matches 848; Conservative 0; Mismatches 384; Indels 3; Gaps 1;
 10 TTAAGAACCTCATACCTGTGTGGCTTTTACTATTTTGGCTTTTATTCATCATCTTTATC 1278

```

Db 202 TTTAAAAAAACCTTAAATGTTTATCCCTTTATTTTGTGATATCTATCTGATTTATCTA 261
QY 70 AATGTTATATCTTTGTTGCTAAAGAGAGCTTGCAATTTATGCTTTTTCGTCGATAGCT 129
Db 262 AATATATATCTTTTGAAC--ATCAACTGTAGGAATTTATGGAGTAAATTAATAAC 318
QY 130 TACCTATTAGTCAAAATGCTCTATCTCTTTTATACAGCCATTTAAGGGAAGGCTGG 189
Db 319 TATCTAGTATCAACTTGGATATCTTTCTTTATGAGCCATTTAAGGAATTCACAT 378
QY 190 CATATAGGTTGAGCCATTTATCCCTTTTATACGAAGATGCTGAGTCATCTGATAG 249
Db 379 GACTATAAAGTTGCTGTTAATCTCTTTTATATGAAGATGCGAGTCAATATTAGAA 438
QY 250 ACCTTAAAGTGTTCAGCAGCAACCTTATCCCTAGCAGAAATTTATGTTGTTGAGAT 309
Db 439 ACCTTAAAGTGTTCAGCAGCAACCTTATCCCTAGCAGAAATTTATGTTGTTGAT 498
QY 310 GGAAGTGTGATGACAGAGTATTAAGCGCATTTGAAGACTATGTCGTGACACTGTGAC 369
Db 499 GGGAGTCAACACAGATGCAATACAAATTAATGAAGATGTAATAGAGATGGAT 558
QY 370 CTATCAAGATGCTATGTTCAATCGGTGAGAAAATCAAGGAAGCGCTCATGACAG 429
Db 559 ATTGTGGAAGCTTATCGTTACCGTTCCTGTCATTAAGGAAGAAACGCCATGCTCAA 618
QY 430 GCTGGGCTTTGAAGATGACAGCTGATGCTTTTTCACCGTTGACTCAGATATAT 489
Db 619 GCGTGGGANTTGAAGATGACGCTGACGTTTTTAAACCGTAGACTCAGATATAT 678
QY 490 ATCTACCCCTGATGCTTTAGAGGATGTTTAAACCTTTAATGACCAACTGTTTTGCT 549
Db 679 ATCTATCCAAATGCTTTAGAGAACTCTTAAGAGTCAATGATGAGACAGTTTATGCT 738
QY 550 GCGAGGCTCAGCTTATGTCAGAAATGAGCAACCAATCTCTTAAACGCTTTGACAGAT 609
Db 739 GCAACAGGACATTTGAATGCTAGAAACAGCAAACTAATCTATTAAACGCACTACAGAT 798
QY 610 ATTCGCTATGATATGCTTTTGGCTTCAAGAGCTGCCAATCCGTTACAGTAATATC 569
Db 799 ATCCGTTAGTATGCTTTGGGTTGAGGCTGCTCAATCAATTAACAGGTAATAT 858
QY 670 CTGTTTGTCTCAGTCCGCTTACGCTTTTACAGACGGAGGTGTTTCTTCAATACATAGAT 729
Db 859 TTAGTTTGTCTCAGACCACTTATGATTTATCGAGTGAAGTGAATTTCTTAACTTAGAG 918
QY 730 AGATACATCAACAGACTCTCTGGTATTCCTGTAAGTATTTGTTGATGACAGGCTGTG 789
Db 919 CGTATTAATAATCAACATCTCCTAGGTTTACCTGTTAGCAATTTGGGATGATGATTTA 978
QY 790 ACCAATATGCACTGATTTAGAAAGACTTTTATCAATCCACTGCTAAATGATATACA 849
Db 979 ACAATATATGATTTAGAGCGACTGTCTACCATCAACAGCTAGATGATGACT 1038
QY 850 GATGTTCTGACAGATGCTACTTACTTGAGCAGCAAAACCGCTGGAACAGTCCCTC 909
Db 1039 GATGATACCTTTCCAAATTAAGAGTTATTTAAGCAACAAATTCGATGCAATATCTTT 1098
QY 910 TTTAGAGAGTCCATTTATCTGTTAAGAAATCATGAAATCTCTTTTGTAGCCCTATGG 969
Db 1099 TTTAGAGATCTATATTTCTGTTTAAANAATCTTTCTAATCCCATCGTTGCTTATGG 1158
QY 970 ACCATCTGAGGTGCTATGTTTATGATGCTTTTATTTCTGTTGTTGATTTCTTTGTA 1029
Db 1159 ACTATTTTGAAGTCTGTTATGTTTATGATGTTGATGTCGCAATTTGGGATCTTTTGT 1218
QY 1030 GGCATGTCAGAGATTTGATGCTCAGGCTTTAGGCTTTTACGCTTTTGTGTTGATTAFTCTCA 1089
Db 1219 AATCAAGCTATTTCAATTAGACCTTATTAACCTTTTGGCTTTTATCCATCATCTTTATC 1278
QY 1090 GTTGGCCGTGTCGGACATTCATTACATGCTTAAGCACCGCTGCTCTCTTCTTGTATCT 1149

```

```

Db 1279 GTTCTTTTATGCTGAATGTTTCATTATATATGTTGGTCAACACTCTGCTAGTTTCTTCT 1338
QY 1150 CCGTTTATGCGGTGCTGCATTTGTTTGCCTACAGCCCTTGAAATTAATATCTCTTTT 1209
Db 1339 COTCTGTATGGAATATTACACTTCTTGTCTTACAGCCCTTAAACTTTTATCTTTATGC 1398
QY 1210 ACTATTAGAATGCTGACTGGGGAACACCTAAAAA 1244
Db 1399 ACCATTAAAAATACGAATGGGAACACGTAAAAA 1433

RESULT 3
US-08-675-499A-1
; Sequence 1, Application US/08675499A
; Patent No. 6492150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwesman, Lundberg, Woessner & Kluth, P. A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2947 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-675-499A-1

Query Match 4.9%; Score 61.6; DB 4; Length 2947;
Best Local Similarity 49.1%; Pred. No. 3.1e-08;
Matches 222; Conservative 0; Mismatches 224; Indels 6; Gaps 2;

QY 474 TGACTCAGATCTATATCTACCTTCATGCTTACAGGAGTGTGTTAAAAACCTT--TAA 530
Db 1140 TGACTCAGATCTATGCTTGACCTTCATCTGTCGATGTTGGATGGTGAAGGCTTTAGAGA 1199
QY 531 TGACCAACTGTTTGTGTCGACGGGTACCTTAATGTCAGAAATAGACAAACCAATCT 590
Db 1200 AGACCCCTGTTGGAGGTGTTGGAGGAGATGCTCCAGATTTTAAACAGATGATGATCTCG 1259
QY 591 CTTAAACAGCTTGACAGATATTCGCTATGATATGCTTTTGGCGTTGAACGAGCTGCCCA 650
Db 1260 GATCTCTCTCCAGCGCTGAGATCTGATGCTGCTTTTAAATATAGAAAGGCGCTGCCA 1319

```

QY 651 ATCCGTACAGTAATATCTGTTGCTCAGTCCGCTTAGCTTTACAGACGCGAGT 710
Db 1320 GCTTATTTTGGTGTGCTCAGTCAATGAGGCTCTCTGGGAATGTACAGAACTCCCT 1379
QY 711 GGTGTTCTTCAATATAGATAGATCAATCAACAGACCTTCTCTGGGTATCTCTGTAAGTAT 770
Db 1380 GCTGATGAATTTGGAAGACTGGTACATCAAGGAATTCATGGGTAAACCAATGCAGTT 1439
QY 771 TGGTATGACAGTGTCTGACCACTATGCACTGATTTAGAAAGACTCTTTATCAATC 830
Db 1440 TGGTACAGAGGACCTTACCAACAGGGTGTAGTCTGGGCTATGCAACTAAATACAC 1499
QY 831 CACTG---CTAATGATTTACAGATGTTCTCTGACAAGATGCTACTTACTTGAAGCAGCA 887
Db 1500 GGTGCTGCTCAAGTCTTACTGAACTCCCATAGATAATCTAGATGGCTGACCAAGCA 1559
QY 888 ARACCGCTGGAAAGTCTCTTTTAGAGAGT 919
Db 1560 GACCGATGGAGCAAGTCTCTACTTCCGAGAGT 1591

RESULT 4

US-09-214-808-1/C
; Sequence 1, Application US/09214808A
; Patent No. 6475793
; GENERAL INFORMATION:
; APPLICANT: Rosenthal, Andre
; APPLICANT: Freiberg, Christoph
; APPLICANT: Perret, Xavier Philippe
; APPLICANT: Broughton, William John
; TITLE OF INVENTION: Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic
; Patent No. 6475793
; TITLE OF INVENTION: Plasmid
; FILE REFERENCE: CARP0068
; CURRENT APPLICATION NUMBER: US/09/214.808A
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: PCR/IB97/00950
; PRIOR FILING DATE: 1997-07-10
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 536165
; TYPE: DNA
; ORGANISM: Rhizobium
US-09-214-808-1

Query Match 4.3%; Score 53.8; DB 4; Length 536165;
Best Local Similarity 46.1%; Pred. No. 7.8e-05;
Matches 181; Conservative 0; Mismatches 212; Indels 0; Gaps 0;
QY 403 AAAATCAAGAAAGCGTCATGCACAGCGCTGGGCTTTGAAAGATCAGACGCTGATGC 462
Db 156991 AAGACGTCGGAAGCGCAAGCGCAGATTGCGGATACGGAATCATCGGAGATTG 156932
QY 463 TTTTTCACCGTGTACATCAGATCTATATCTACCTGATGCTTTAGAGGAGTTGTAATA 522
Db 156931 GTGCTACAGCTGTACGCGACAGCAACATTCGCCGAGCTAGTCAGAAATTCGCCCTG 156872
QY 523 ACCTTTAATGACCAACTGTTTGTGTCGCGAGCGGTTCACCTTATGTCAGAAATAGACAA 582
Db 156871 AAGATGTACAGTCCCGCGTGGCGCGCGATGGGTGAGTGTAGCGCGACAGCAGCAGC 156812
QY 583 ACCAATCTCTTACAGCGTTGACAGATATTCGCTATGATAATGCTTTTGGCGTTGACAGCA 642
Db 156811 GACATATGGCTACGCGGTTGATGACATGAGTACTGGCTGCTGCAACAGAGAGCG 156752
QY 643 GCTGCCAATCCGTTACAGTAATATCTTTGTCAGTCGCGCTTACGCTTTACAGA 702
Db 156751 GCAGCAGGCTCGCTTTTGAGCGGCTTATGTTGTCGCGCGCTGTCATGCGG 156692
QY 703 CCGAGGTGTTGTTCTTCAACATAGATAGATACATCAACAGAGCTTCTCGGATTCCT 762
Db 156692 CCGAGGTGTTGTTCTTCAACATAGATAGATACATCAACAGAGCTTCTCGGATTCCT 762

Db 156691 CGTCCCGACTCTCTATTGCTGCTGAATAAATACAGAGCAAGCAACTGTTTCGAGGACGCA 156632
QY 763 GTAAATGTTGGTATGATGACAGGTCGCTTGACCAAC 795
Db 156631 AGCGACTTCGGGGAAGACCGCCACCTCACAAATC 156599
RESULT 5
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-Fls
US-08-232-463-14
Query Match 3.6%; Score 45.6; DB 1; Length 7218;
Best Local Similarity 3.4%; Pred. No. 0.0026;
Matches 12; Conservative 197; Mismatches 141; Indels 0; Gaps 0;
QY 904 TCCTCTTITAGAGATCCATATTTCTGTTAAGAAATCATGAACAATCCTTTTGTAGCC 963
Db 1101 YY 1160
QY 964 CTATGACCACTTATGAGTGCTATGTTTANGATCTGTTTATCTGTGPGGATTC 1023
Db 1161 YY 1220
QY 1024 TTGTAGGCAATGTCAGAGATTTGATGGCTCAGGTTTAGCCTTCTGTTGATATC 1083
Db 1221 YY 1280
QY 1084 TTCATTGTCGCTGTGCGAACAATCATACATGCTTAAGCAGCGCTGCTCTTG 1143

[illegible]

```

1  RESULT 6
2  US-08-036-210-6
3  ; Sequence 6, Application US/08036210
4  ; Patent No. 5585233
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Moller, Niels P.H.
7  ; APPLICANT: Moller, Karin B.
8  ; APPLICANT: Ullrich, Axel
9  ; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE
10 ; TITLE OF INVENTION: PHOSPHATASE
11 ; NUMBER OF SEQUENCES: 45
12 ; CORRESPONDENCE ADDRESS:
13 ; ADDRESSEE: PENNIE & EDMONDS
14 ; STREET: 1155 Avenue of the Americas
15 ; CITY: New York
16 ; STATE: New York
17 ; COUNTRY: U.S.A.
18 ; ZIP: 10036-2711
19 ; COMPUTER READABLE FORM:
20 ; MEDIUM TYPE: Floppy disk
21 ; COMPUTER: IBM PC compatible
22 ; OPERATING SYSTEM: PC-DOS/MS-DOS
23 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
24 ; CURRENT APPLICATION DATA:
25 ; APPLICATION NUMBER: US/08/036,210
26 ; FILING DATE: 23-MAR-1993
27 ; CLASSIFICATION: 435
28 ; ATTORNEY/AGENT INFORMATION:
29 ; NAME: Mistrock, S. Leslie
30 ; REGISTRATION NUMBER: 18,872
31 ; REFERENCE/DOCKET NUMBER: 7683-025
32 ; TELECOMMUNICATION INFORMATION:
33 ; TELEPHONE: 212-790-9090
34 ; TELEFAX: 212-869-8864/9741
35 ; TELEX: 66141 PENNIE
36 ; INFORMATION FOR SEQ ID NO: 6:
37 ; SEQUENCE CHARACTERISTICS:
38 ; LENGTH: 2173 base pairs
39 ; TYPE: nucleic acid
40 ; STRANDEDNESS: double
41 ; TOPOLOGY: unknown
42 ; MOLECULE TYPE: cDNA
43 ; FEATURE:
44 ; NAME/KEY: misc.feature
45 ; LOCATION: 1630
46 ; OTHER INFORMATION: /note= "N=X-unknown nucleotide"
47 US-08-036-210-6

```

	Query Match	3.0%;	Score 37.8;	DB 1;	Length 2173;
	Best Local Similarity	48.4%;	Pred. No. 0.29;		
	Matches 105;	Conservative 0;	Mismatches 112;	Indels 0;	Gaps 0;
QY	874	TACITGAAGCAGCAAAACCGCTGGACATGTCCTTTTAGAGAGTCCATATATTTCCTTT	933		
Db	1877	TACTGTGTGAAGTCAAAACAAAGAAATTAACCATCTATAGCAAAATTCAGGTTTCTTTA	1936		
QY	934	AAGAAATCATGAACAATCTTTTGTAGCCCTATGTGACCATACITGAGTGCTCATGTTT	993		
Db	1937	TASAAATCTTTCAGCCCTCCATCTTATTAATAGTGCATATGTGTAAGTTTGAATAT	1996		
QY	994	ATGATGCTGTTTATTCTGTGTGGTATTTCTTTGTAGGCAATGTCAGAGAAATTCATTGG	1053		

```

Db      1997 ATGAACACTATTGTCATAGATTTCAATAAAGAGTAATAATAGTATTATGTCTT 2056
QY      1054 CTCAGGGTTTTAGCCCTTTCTGCGTGATTAATCTTCATTG 1090
DB      2057 CTATGATGAAGATATATCTTATGCTTATTTCCGCIG 2093

RESULT 7
US-08-449-609-6
; Sequence 6, Application US/08449609
; Patent No. 5952212
; GENERAL INFORMATION:
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Moller, Karin B.
; APPLICANT: Ulirich, Axel
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FatenIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,609
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/036,210
; FILING DATE: 23-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2173 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1630
; OTHER INFORMATION: /note= "N-X-unknown nucleotide"
US-08-449-609-6

```

		Query Match 3.0%;	Score 37.8;	DB 2;	Length 2173;
		Best Local Similarity 48.4%;	Pred. No. 0.29;		
		Matches 105;	Conservative 0;	Mismatches 112;	Indels 0; Gaps 0;
QY	874	TACTTGAGCAGCAAAACCGCTGGACAACAGTCCTCTTTTAGAGAGTCATTATTTCTGTT	933		
Db	1877	TACTTGGTGAACCTGAAACAAACAGAAATTAACCATCTATAGCAAACTCAGAGTTTCCTTA	1936		
QY	934	AGAAATATGAAACAATCCTTTTGTAGCCCTATGGACCATCTTGCAGGTGCTCATGTTT	993		
Db	1937	TAGAAATCTTTTCAGCCCTCCATCTATTAAATAGTGACATGTGGTAGTTTTCGAATTAT	1996		
QY	994	ATGATGCTTTTATTCTGTGTGGTGAATTCCTTTGTAGGCCAATGTCAGAGAATTTGATTGG	1053		

Db 1997 ATGAACATCTTTTCATAGATTTCAATTAAGAGTAATAAATAGATTAATTAATGCTCTT 2056
QY 1054 CTCAGGTTTACGCTTCTGCTGATTAATTCATG 1090
Db 2057 CTATGATAAGAGTATATCTTAATGCTTATTCGCTG 2093

RESULT 8

US-09-361-096A-6
; Sequence 6, Application US/09361096A
; Patent No. 6492495
; GENERAL INFORMATION:
; APPLICANT: MOLLER, NIELS P.H.
; APPLICANT: MOLLER, KARIN B.
; APPLICANT: ULLRICH, AXEL
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE PHOSPHATASE
; FILE REFERENCE: 038602/0686
; CURRENT APPLICATION NUMBER: US/09/361.096A
; CURRENT FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 08/449,609
; PRIOR FILING DATE: 1995-05-24
; PRIOR APPLICATION NUMBER: 08/036,210
; PRIOR FILING DATE: 1995-03-23
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2173
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PTP-S31C
; NAME/KEY: modified base
; LOCATION: (1)..(1630)
; OTHER INFORMATION: n - unknown nucleotide
; NAME/KEY: CDS
; LOCATION: (65)..(1066)
US-09-361-096A-6

Query Match 3.0%; Score 37.8; DB 4; Length 2173;
Best Local Similarity 48.4%; Pred. No. 0.29;
Matches 105; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
QY 874 TACTGAGCAGCAAAACCGCTGGAAAGTCTCTTTAGAGAGTCCATTAATTTCTGTT 933
Db 1877 TACTGTTGGAAGTGAACCAAGAAATTAACCATCTATAGCAATTCAGGTTTCTTA 1936
QY 934 AAGAAATCATGAAACATCCCTTTGTAGCCCTATGGACCATCTAGGTGCTTAATGTT 993
Db 1937 TAGAAATCTTTTCAGCCTCCATCTTAATAATAGTGAATGTTGAATTTTGAATAT 1996
QY 994 ATGATGCTTTTATCTGTGCTGCTGATTTCTTTGTAGCAATGTCAGAGAAATTTGATGG 1053
Db 1997 ATGAACATCTTTTGTCATAGATTTCAATTAAGAGTAATAAATAGTATTAATTAATGCTCTT 2056
QY 1054 CTCAGGTTTACGCTTCTGCTGATTAATTCATG 1090
Db 2057 CTATGATAAGAGTATATCTTAATGCTTATTCGCTG 2093

RESULT 9

US-08-036-210-10
; Sequence 10, Application US/08036210
; Patent No. 5585233
; GENERAL INFORMATION:
; APPLICANT: MOLLER, NIELS P.H.
; APPLICANT: MOLLER, KARIN B.
; APPLICANT: ULLRICH, AXEL
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas

CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/036,210
FILING DATE: 23-MAR-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7683-025
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864/9741
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2309 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1766
OTHER INFORMATION: /note= "N-x-unknown nucleotide"
US-08-036-210-10

Query Match 3.0%; Score 37.8; DB 1; Length 2309;
Best Local Similarity 48.4%; Pred. No. 0.3;
Matches 105; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
QY 874 TACTGAGCAGCAAAACCGCTGGAAAGTCTCTTTAGAGAGTCCATTAATTTCTGTT 933
Db 2013 TACTGTTGGAAGTGAACCAAGAAATTAACCATCTATAGCAATTCAGGTTTCTTA 2072
QY 934 AAGAAATCATGAAACATCCCTTTGTAGCCCTATGGACCATCTAGGTGCTTAATGTT 993
Db 2073 TAGAAATCTTTTCAGCCTCCATCTTAATAATAGTGAATGTTGAATTTTGAATAT 2132
QY 994 ATGATGCTTTTATCTGTGCTGCTGATTTCTTTGTAGCAATGTCAGAGAAATTTGATGG 1053
Db 2133 ATGAACATCTTTTGTCATAGATTTCAATTAAGAGTAATAAATAGTATTAATTAATGCTCTT 2192
QY 1054 CTCAGGTTTACGCTTCTGCTGATTAATTCATG 1090
Db 2193 CTATGATAAGAGTATATCTTAATGCTTATTCGCTG 2229

RESULT 10

US-08-449-609-10
; Sequence 10, Application US/08449609
; Patent No. 5952212
; GENERAL INFORMATION:
; APPLICANT: MOLLER, NIELS P.H.
; APPLICANT: MOLLER, KARIN B.
; APPLICANT: ULLRICH, AXEL
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,609
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/036,210
; FILING DATE: 23-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2309 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1766
; OTHER INFORMATION: /note= "N=x-unknown nucleotide"
;
US-08-449-609-10

Query Match          3.0%; Score 37.8; DB 2; Length 2309;
Best Local Similarity 48.4%; Pred. No. 0.3;
Matches 105; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 874 TACTTGAAGCAGCAAAACCGCTGGAAACGCTCTTTTGTAGAGAGTCCATTATTTCTGTT 933
Db 2013 TACTTGGTGAAGCTGAAACAAAGAAATTAACCATCTATAGCAAAATTCAGGTTTCTTTA 2072

QY 934 AGAAATCATGACAAATCCCTTTTGTAGCCCTTAGGACCATACTTGGAGTGCTCTATGTT 993
Db 2073 TAGAAATCTTTCAGCCCTCCATCTTATTAATAGTGAATGTGTAAGTTTGAATTAT 2132

QY 994 ATGATGCTGTTTATCTGTGGTGGATTTCTTTGTAGCAATGTCAGAGAAATTTGATTGG 1053
Db 2133 ATGACTCATTTTGTATAGATTTCAATTAAGAGTAATAAATAGTATTAATTAATGCTCTT 2192

QY 1054 CTCAGGGTTTAGCCCTTTCGTGGTGAATTAATCTTCATTG 1090
Db 2193 CTATGATAAGAAGTATATCTTATGCTTATTTCCGCTG 2229

RESULT 11
US-09-361-096A-10
; Sequence 10, Application US/09361096A
; Patent No. 6492495
; GENERAL INFORMATION:
; APPLICANT: MOLLER, NIELS P.H.
; APPLICANT: MOLLER, KARIN B.
; APPLICANT: ULLRICH, AXEL
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE PHOSPHATASE
; FILE REFERENCE: 038602/0686
; CURRENT APPLICATION NUMBER: US/09/361,096A
; CURRENT FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 08/449,609
; PRIOR FILING DATE: 1995-05-24
; PRIOR APPLICATION NUMBER: 08/036,210
; PRIOR FILING DATE: 1995-03-23
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,609
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/036,210
; FILING DATE: 23-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2309 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1766
; OTHER INFORMATION: /note= "N=x-unknown nucleotide"
;
US-08-449-609-10

Query Match          3.0%; Score 37.8; DB 4; Length 2309;
Best Local Similarity 48.4%; Pred. No. 0.3;
Matches 105; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 874 TACTTGAAGCAGCAAAACCGCTGGAAACGCTCTTTTGTAGAGAGTCCATTATTTCTGTT 933
Db 2013 TACTTGGTGAAGCTGAAACAAAGAAATTAACCATCTATAGCAAAATTCAGGTTTCTTTA 2072

QY 934 AGAAATCATGACAAATCCCTTTTGTAGCCCTTAGGACCATACTTGGAGTGCTCTATGTT 993
Db 2073 TAGAAATCTTTCAGCCCTCCATCTTATTAATAGTGAATGTGTAAGTTTGAATTAT 2132

QY 994 ATGATGCTGTTTATCTGTGGTGGATTTCTTTGTAGCAATGTCAGAGAAATTTGATTGG 1053
Db 2133 ATGACTCATTTTGTATAGATTTCAATTAAGAGTAATAAATAGTATTAATTAATGCTCTT 2192

QY 1054 CTCAGGGTTTAGCCCTTTCGTGGTGAATTAATCTTCATTG 1090
Db 2193 CTATGATAAGAAGTATATCTTATGCTTATTTCCGCTG 2229

RESULT 12
US-08-036-210-14
; Sequence 14, Application US/08036210
; Patent No. 5585233
; GENERAL INFORMATION:
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Moller, Karin B.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE PHOSPHATASE
; FILE REFERENCE: 038602/0686
; CURRENT APPLICATION NUMBER: US/08/036,210
; CURRENT FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 08/449,609
; PRIOR FILING DATE: 1995-05-24
; PRIOR APPLICATION NUMBER: 08/036,210
; PRIOR FILING DATE: 1995-03-23
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/036,210
; FILING DATE: 23-MAR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2692 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```

```
;
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2149
; OTHER INFORMATION: /note= "N-x-unknown nucleotide"
US-08-036-210-14

Query Match          3.0%; Score 37.8; DB 1; Length 2692;
Best Local Similarity 48.4%; Pred. No. 0.32;
Matches 105; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 874 TACTTGAAGCAGCAAAACCGCTGGAACAGTCCTCTTTAGAGAGTCCTATTTCTGTT 933
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2396 TACTTGTGGAACTGAAACAAGAAATTAACCATCTATAGCAAAATTCAGGTTTCTTTA 2455
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 934 AAGAAATCATGAACAATCCTTTTGTAGCCCTATGGACCATCTTGGAGTGTCTATGTTT 993
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2456 TAGAAATCTTTTCAGCTCCATCTTATTAAATAGTGACAATGTGTAAGTTTGAATTAT 2515
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 994 ATGATGCTTGTATTCTGTGGTGGATTTCTTTGTAGGCAATGTCAGAGAATTTGATGG 1053
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2516 ATGAATCATTTTGTGCATAGATTTCAATTAAGAGTAATAATAGTATTATATGCTCTT 2575
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1054 CTCAGGTTTTCAGCTTCTGTGATTAATCTTCATTG 1090
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2576 CTATGATAAGAGTATATCTTATGCTTATTCGCTG 2612
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-08-449-609-14
; Sequence 14, Application US/08449609
; Patent No. 5952212
; GENERAL INFORMATION:
; APPLICANT: Moller, Niels P.H.
; APPLICANT: Moller, Karin B.
; APPLICANT: Ullrich, Axel
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE
; TITLE OF INVENTION: PHOSPHATASE
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,609
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/036,210
; FILING DATE: 23-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7683-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2692 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown

;
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2149
; OTHER INFORMATION: /note= "N-x-unknown nucleotide"
US-08-449-609-14

Query Match          3.0%; Score 37.8; DB 2; Length 2692;
Best Local Similarity 48.4%; Pred. No. 0.32;
Matches 105; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 874 TACTTGAAGCAGCAAAACCGCTGGAACAGTCCTCTTTAGAGAGTCCTATTTCTGTT 933
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2396 TACTTGTGGAACTGAAACAAGAAATTAACCATCTATAGCAAAATTCAGGTTTCTTTA 2455
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 934 AAGAAATCATGAACAATCCTTTTGTAGCCCTATGGACCATCTTGGAGTGTCTATGTTT 993
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2456 TAGAAATCTTTTCAGCTCCATCTTATTAAATAGTGACAATGTGTAAGTTTGAATTAT 2515
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 994 ATGATGCTTGTATTCTGTGGTGGATTTCTTTGTAGGCAATGTCAGAGAATTTGATGG 1053
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2516 ATGAATCATTTTGTGCATAGATTTCAATTAAGAGTAATAATAGTATTATATGCTCTT 2575
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1054 CTCAGGTTTTCAGCTTCTGTGATTAATCTTCATTG 1090
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2576 CTATGATAAGAGTATATCTTATGCTTATTCGCTG 2612
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
US-09-361-096A-14
; Sequence 14, Application US/09361096A
; Patent No. 6492495
; GENERAL INFORMATION:
; APPLICANT: MOLLER, NIELS P.H.
; APPLICANT: MOLLER, KARIN B.
; APPLICANT: ULLRICH, AXEL
; TITLE OF INVENTION: PTP-S31: A NOVEL PROTEIN TYROSINE PHOSPHATASE
; FILE REFERENCE: 038602/0686
; CURRENT APPLICATION NUMBER: US/09/361,096A
; CURRENT FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 08/449,609
; PRIOR FILING DATE: 1995-05-24
; PRIOR APPLICATION NUMBER: 08/036,210
; PRIOR FILING DATE: 1995-03-23
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 14
; LENGTH: 2692
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1)..(2692)
; OTHER INFORMATION: n = unknown nucleotide
; NAME/KEY: CDS
; LOCATION: (92..139, 259..1414)
US-09-361-096A-14

Query Match          3.0%; Score 37.8; DB 4; Length 2692;
Best Local Similarity 48.4%; Pred. No. 0.32;
Matches 105; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 874 TACTTGAAGCAGCAAAACCGCTGGAACAGTCCTCTTTAGAGAGTCCTATTTCTGTT 933
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2396 TACTTGTGGAACTGAAACAAGAAATTAACCATCTATAGCAAAATTCAGGTTTCTTTA 2455
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 934 AAGAAATCATGAACAATCCTTTTGTAGCCCTATGGACCATCTTGGAGTGTCTATGTTT 993
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2456 TAGAAATCTTTTCAGCTCCATCTTATTAAATAGTGACAATGTGTAAGTTTGAATTAT 2515
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 994 ATGATGCTTGTATTCTGTGGTGGATTTCTTTGTAGGCAATGTCAGAGAATTTGATGG 1053
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2516 ATGAATCATTTTGTGCATAGATTTCAATTAAGAGTAATAATAGTATTATATGCTCTT 2575
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: October 2, 2003, 09:03:35
Job time : 86 secs

RESULT 15
US-08-036-210-21

GENERAL INFORMATION:

APPLICANT: Moller

APPLICANT: Moller

APPLICANT: Ullrich

TITLE OF INVENTION

NUMBER OF SEQUENCE

ADDRESSEE: PENN

CITY: New York

COUNTRY: U.S.A.

COMPUTER READABLE

COMPUTER: IBM P

SOFTWARE: Paten

APPLICATION NUMB

CLASSIFICATION:

NAME: Misrock,

REFERENCE/DOCKET

TELEPHONE: 212-

TELEX: 66141 PE

SEQUENCE CHARACTER

TYPE: nucleic acid

TOPOLOGY: unknc

FEATURE:

LOCATION: 3430

08-036-210-21

Query Match

Matches 105; Conser

874 TACTGAAG

3677 TACTTGGTC

934 AAGAAATC

3131 TAGAAAAATC

994 ATGATGCTT
1111

379/ ATGAATCTC

5COT 4CAGGT

385 / C1ATGATAT

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2003, 09:10:12 ; Search time 345 Seconds
(without alignments)
9271.834 Million cell updates/sec

Title: US-09-469-200D-1
Perfect score: 1254
Sequence: 1 atgagacattataaacct.....caogtaaaattattataa 1254

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1708419 seqs, 1275431651 residues

Total number of hits satisfying chosen parameters: 3416838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	ID	Description
1	1254	100.0	1254	10	US-09-879-959-1	Sequence 1, Appli
2	1254	100.0	1254	14	US-10-011-768B-1	Sequence 1, Appli
3	1254	100.0	1254	14	US-10-011-771B-1	Sequence 1, Appli
4	1254	100.0	1254	14	US-10-172-527-1	Sequence 1, Appli
5	1225.4	97.7	1251	12	US-10-326-185-1	Sequence 1, Appli
6	1152	91.9	5158	12	US-10-326-185-108	Sequence 108, App
7	623.2	49.7	1251	12	US-10-326-185-102	Sequence 102, App
8	623.2	49.7	3466	14	US-10-172-527-11	Sequence 11, Appl
9	604.6	48.2	1440	14	US-10-172-527-13	Sequence 13, Appl
10	604.6	48.2	1512	14	US-10-124-222-1	Sequence 1, Appli
11	604.6	48.2	1512	14	US-10-117-795-1	Sequence 1, Appli
12	601.4	48.0	1257	12	US-10-326-185-92	Sequence 92, Appl
13	282.8	22.6	1200	14	US-10-172-527-17	Sequence 17, Appl
14	61.6	4.9	1659	11	US-09-902-939-2	Sequence 2, Appli
15	61.6	4.9	4194	14	US-10-262-526-3	Sequence 3, Appli
16	54	4.3	54	12	US-10-326-185-39	Sequence 39, Appl

c 17 53.8 4.3 536165 11 US-09-939-964-1 Sequence 1, Appli
18 49.4 3.9 1665 11 US-09-902-939-3 Sequence 3, Appli
19 47.2 3.8 3003 11 US-09-918-624B-64 Sequence 64, Appli
20 47.2 3.8 3003 14 US-10-262-526-1 Sequence 1, Appli
21 45.6 3.6 4018 12 US-09-814-353-19942 Sequence 19942, A
22 43 3.4 1752 11 US-09-902-939-1 Sequence 1, Appli
23 43 3.4 3387 14 US-10-084-817-184 Sequence 184, App
24 43 3.4 7488 12 US-10-311-455-1085 Sequence 1085, Ap
25 43 3.4 9899 12 US-10-311-455-2416 Sequence 2416, Ap
26 41.6 3.3 1740 10 US-09-879-959-8 Sequence 8, Appli
27 41.6 3.3 1740 14 US-10-011-768B-7 Sequence 7, Appli
28 41.6 3.3 1740 14 US-10-011-771B-7 Sequence 7, Appli
29 41.6 3.3 1740 14 US-10-172-527-8 Sequence 8, Appli
30 41 3.3 6301 12 US-10-311-455-26 Sequence 26, Appli
31 39 3.1 3528 10 US-09-070-927A-189 Sequence 189, App
32 38.8 3.1 888 12 US-10-032-585-6433 Sequence 6433, Ap
33 38.8 3.1 9817 12 US-10-311-455-1341 Sequence 1341, Ap
34 38 3.0 6485 12 US-10-311-455-1781 Sequence 1781, Ap
35 38 3.0 7049 12 US-10-311-455-130 Sequence 130, App
36 38 3.0 7049 12 US-10-240-452-6 Sequence 6, Appli
37 38 3.0 7119 12 US-10-311-455-266 Sequence 266, App
38 38 3.0 15306 12 US-10-240-453-184 Sequence 184, App
39 38 3.0 15306 14 US-10-239-676-162 Sequence 162, App
40 37.8 3.0 2173 12 US-10-314-232-6 Sequence 6, Appli
41 37.8 3.0 2309 12 US-10-314-232-10 Sequence 10, Appli
42 37.8 3.0 2692 12 US-10-314-232-14 Sequence 14, Appli
43 37.8 3.0 3973 12 US-10-314-232-21 Sequence 21, Appli
44 37.6 3.0 9243 12 US-10-240-453-127 Sequence 127, App
45 37.4 3.0 600 13 US-10-027-632-84036 Sequence 84036, A

ALIGNMENTS

RESULT 1

US-09-879-959-1
; Sequence 1, Application US/09879959
; Patent No. US20020160489A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H
; APPLICANT: Kumari, Kshama
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: STREPTOCOCCUS EQUISIMILIS HYALURONAN SYNTHASE GENE AND EXPRESS
; FILE REFERENCE: 3554 049
; CURRENT APPLICATION NUMBER: US/09/879,959
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 09/469,200
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1254
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-879-959-1

Query Match 100.0%; Score 1254; DB 10; Length 1254;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGACATTATAAACCTCATACTGTGCGCCCTTAGTATTTTGGTACGTTTG 60
|||||
Db 1 ATGAGACATTATAAACCTCATACTGTGCGCCCTTAGTATTTTGGTACGTTTG 60
|||||
QY 61 ATTACGCTCAATGTTATCTCTTTGGTGTAAAGAGCTTGTCAATTATGCGCTTTTG 120
|||||
Db 61 ATTACGCTCAATGTTATCTCTTTGGTGTAAAGAGCTTGTCAATTATGCGCTTTTG 120
|||||
QY 121 CTGACGCTTACTATTAGTCAAAATGCTTCCTTTTTCACAGCCATTAAGGA 180
|||||

121	CTGATAGCTTACCTATTATAGTCAAAATGCTCTTATTCCTTTTTTTTACAGCCATTAAAGGA	180	
QY	181	AGGCTGGGCAATATAAGTGTGCACCAATTATTCCTCTTTATAACAGAGATGCTGAGTCA	240
Db	181	AGGCTGGGCAATATAAGTGTGCACCAATTATTCCTCTTTATAACAGAGATGCTGAGTCA	240
QY	241	TTGCTAGACGCTTAAAAAGTGTTCAGCAGCAAACTATCCCTAGCAGAAATTTATGTT	300
Db	241	TTGCTAGACGCTTAAAAAGTGTTCAGCAGCAAACTATCCCTAGCAGAAATTTATGTT	300
QY	301	GTTCAGGATGGAAGTGTGATGAGACGAGTATTAAAGCGATTGAAGACTATGTCCGTGAC	360
Db	301	GTTCAGGATGGAAGTGTGATGAGACGAGTATTAAAGCGATTGAAGACTATGTCCGTGAC	360
QY	361	ACTGTGTGACCTATCAAGCAATGTCATGTTTCATCGTGCAGAGAAAAAATCAAGAAAGCGT	420
Db	361	ACTGTGTGACCTATCAAGCAATGTCATGTTTCATCGTGCAGAGAAAAAATCAAGAAAGCGT	420
QY	421	CATGCACAGGCTGGGCTTTGAAAGATCAGACGCTGATGCTTTTTGACCGTGTGACTCA	480
Db	421	CATGCACAGGCTGGGCTTTGAAAGATCAGACGCTGATGCTTTTTGACCGTGTGACTCA	480
QY	481	GATACTTATATCCTACCTGATGCTTTAGAGAGTTGTTAAAAACCTTTTAAATGACCCAACT	540
Db	481	GATACTTATATCCTACCTGATGCTTTAGAGAGTTGTTAAAAACCTTTTAAATGACCCAACT	540
QY	541	GTTTTGTCTGCACGGGTCACTTAAATGTCAGAATAAGACAAACCAATCTCTTAACACGC	600
Db	541	GTTTTGTCTGCACGGGTCACTTAAATGTCAGAATAAGACAAACCAATCTCTTAACACGC	600
QY	601	TTGACAGATATTCGCTATGATAATGCTTTTGGCGTTGAAGAGCTGCCCAATCCGTTACA	660
Db	601	TTGACAGATATTCGCTATGATAATGCTTTTGGCGTTGAAGAGCTGCCCAATCCGTTACA	660
QY	661	GSTAAATCTCTGTTTGCTCAGGTCGCGCTTACGTTTACAGACGGAGGCTGTGTCCT	720
Db	661	GSTAAATCTCTGTTTGCTCAGGTCGCGCTTACGTTTACAGACGGAGGCTGTGTCCT	720
QY	721	AACATAGATAGATACATCAACACAGACTTCCTGGGTATTCCTGTAAGTATTTGGTGATGAC	780
Db	721	AACATAGATAGATACATCAACACAGACTTCCTGGGTATTCCTGTAAGTATTTGGTGATGAC	780
QY	781	AGGTCTGTGACCACTATGCACTGATTTAGGAAGACTGTTTATCAATCCACTGCTAAA	840
Db	781	AGGTCTGTGACCACTATGCACTGATTTAGGAAGACTGTTTATCAATCCACTGCTAAA	840
QY	841	TGTAATACAGATTCCTCGACACATGCTACTTACTTGAAGCAGCAAAACCGTGGAAAC	900
Db	841	TGTAATACAGATTCCTCGACACATGCTACTTACTTGAAGCAGCAAAACCGTGGAAAC	900
QY	901	AAAGTCTCTTTTAGAGATGCAATTTCTGTTTAAAGAAATCATGAACAAATCCCTTTGTA	960
Db	901	AAAGTCTCTCTTTTAGAGATGCAATTTCTGTTTAAAGAAATCATGAACAAATCCCTTTGTA	960
QY	961	GCCTATGGAACACTTGAAGTGTCTATGTTTATGATGCTTTGTTATTCGTGGTGGAT	1020
Db	961	GCCTATGGAACACTTGAAGTGTCTATGTTTATGATGCTTTGTTATTCGTGGTGGAT	1020
QY	1021	TTCTTTGTAGGCAATGTCAGAGAAATTTGATTTGGCTCAGGGTTTTAGCCCTTCGTGGTAT	1080
Db	1021	TTCTTTGTAGGCAATGTCAGAGAAATTTGATTTGGCTCAGGGTTTTAGCCCTTCGTGGTAT	1080
QY	1081	ATCTTCATTTGCTGCCCTGTGTCGGAACATTCATACATGCTTAAGCACCCGCTGCCTTC	1140
Db	1081	ATCTTCATTTGCTGCCCTGTGTCGGAACATTCATACATGCTTAAGCACCCGCTGCCTTC	1140
QY	1141	TTGTTATCTCCGTTTTATGGGTCGCTGCATTTGTTTGTTCCTACAGCCCTTGAATATAT	1200
Db	1141	TTGTTATCTCCGTTTTATGGGTCGCTGCATTTGTTTGTTCCTACAGCCCTTGAATATAT	1200
QY	1201	TCCTTTTTACTATTAGAAATGCTGACTGGGGACACGCTAAAAAATATATATA	1254
Db	1201	TCCTTTTTACTATTAGAAATGCTGACTGGGGACACGCTAAAAAATATATATA	1254

```

RESULT 2
US-10-011-768B-1
; Sequence 1, Application US/10011768B
; Publication No. US2003007322A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H.
; APPLICANT: DeAngelis, Paul
; APPLICANT: Kumari, Kshama
; TITLE OF INVENTION: Hyaluronan Synthase Gene and Uses Thereof
; FILE REFERENCE: 3554.011
; CURRENT APPLICATION NUMBER: US/10/011,768B
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 09/178,851
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: US 60/064,435
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1254
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-10-011-768B-1

```

Query Match	100.0%;	Score 1254;	DB 14;	Length 1254;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1254;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGAGAACATTAAAAACCTCATACTGTTGGTGGCCCTTTAGTATTTTTTGGGTACTGTGG	60	
DB				
QY	1	ATGAGAACATTAAAAACCTCAAACTCTGTGGCCCTTAGTATTTTTTGGGTACTGTGG	60	
DB				
QY	61	ATTTAGCGTCAATGTTATCTCTTTGGTGCTAAAGAGACCTGTGCAATTTATGGCTTTTGG	120	
DB				
QY	61	ATTTAGCGTCAATGTTATCTCTTTGGTGCTAAAGAGACCTGTGCAATTTATGGCTTTTGG	120	
DB				
QY	121	CTGNTAGCTTACCTATTAGTCAAAATGTCCTTATCCTTTTTTACAAGCCATTTAAGGGA	180	
DB				
QY	121	CTGATAGCTTACCTATTAGTCAAAATGTCCTTATCCTTTTTTACAAGCCATTTAAGGGA	180	
DB				
QY	181	AGGGTGGGCAATATAAGSTTGCAGCCATTATTCCTCTTATACGAGAGTCTGTGATCA	240	
DB				
QY	181	AGGGTGGGCAATATAAGSTTGCAGCCATTATTCCTCTTATACGAGAGTCTGTGATCA	240	
DB				
QY	241	TTGCTAGACACCTTAAAGAGTGTTTCAGCAGCAAACTATCCCTACGACAAATTTATGTT	300	
DB				
QY	241	TTGCTAGACACCTTAAAGAGTGTTTCAGCAGCAAACTATCCCTACGACAAATTTATGTT	300	
DB				
QY	301	GTTGACGATGGAAGTGCTGATGAGACAGGTATTAAAGCCATTGAAGACATATGTCGCTGAC	360	
DB				
QY	361	ACTGGTGACCTATCAAGCAATGTCAATGTTTCATCGTCAAGAGAAAATCAGGAAAGCGT	420	
DB				
QY	361	ACTGGTGACCTATCAAGCAATGTCAATGTTTCATCGTCAAGAGAAAATCAGGAAAGCGT	420	
DB				
QY	421	CATGCACAGGCGTGGCGCTTTGAAAGATCAGACGCTGATGCTTTTTGACCGTGTGATCA	480	
DB				
QY	421	CATGCACAGGCGTGGCGCTTTGAAAGATCAGACGCTGATGCTTTTTGACCGTGTGATCA	480	
DB				
QY	481	GATACATTATATCTACCTGATGCTTTAGAGAGGTGTTTAAAAACCTTTAATGACCCCACT	540	
DB				
QY	481	GATACATTATATCTACCTGATGCTTTAGAGAGGTGTTTAAAAACCTTTAATGACCCCACT	540	
DB				
QY	541	GTTTTTGGTGGCAGGGGTCACTTAAATGTCAGAAATAGACAAACCAATCTCTTTAACACGC	600	
DB				
QY	541	GTTTTTGGTGGCAGGGGTCACTTAAATGTCAGAAATAGACAAACCAATCTCTTTAACACGC	600	
DB				
QY	601	TTGACAGATATTGCGCTATGATATATGCTTTTTGGCGTTGAACAGAGTGCCCAATCCGTTTACA	660	
DB				
QY	601	TTGACAGATATTGCGCTATGATATATGCTTTTTGGCGTTGAACAGAGTGCCCAATCCGTTTACA	660	
DB				

QY 661 GGTAAATATCTTTGTTGCTCAGGTCGCTTAGCTTTACAGACGCGAGGTGTTTCCT 720
 DB 661 GGTAAATATCTTTGTTGCTCAGGTCGCTTAGCTTTACAGACGCGAGGTGTTTCCT 720
 QY 721 AACATAGATAGATACATCAACAGACGCTTCCCTGGGTATTCCTGTAAGTATGGTGATGAC 780
 DB 721 AACATAGATAGATACATCAACAGACGCTTCCCTGGGTATTCCTGTAAGTATGGTGATGAC 780
 QY 781 AGGTGCTTGACCACTATGCAACTGATTAGGAAGAGCTGTTTATCAATCCACTGCTAAA 840
 DB 781 AGGTGCTTGACCACTATGCAACTGATTAGGAAGAGCTGTTTATCAATCCACTGCTAAA 840
 QY 841 TGTAATACAGATGTTCCCTGACAAGATGCTTACTTCTGTAAGAGAGAGCTGTTTATCAATCCACTGCTAAA 900
 DB 841 TGTAATACAGATGTTCCCTGACAAGATGCTTACTTCTGTAAGAGAGAGCTGTTTATCAATCCACTGCTAAA 900
 QY 901 AGTCCCTTTTAGAGAGTCCATTAATTCGTTAGGAAGATCATGAACAAATCCTTTGTA 960
 DB 901 AGTCCCTTTTAGAGAGTCCATTAATTCGTTAGGAAGATCATGAACAAATCCTTTGTA 960
 QY 961 GGCCTATGAGCACTACTGAGGTGCTATGTTTATGANGCTTGTATTCCTGCTGGAT 1020
 DB 961 GGCCTATGAGCACTACTGAGGTGCTATGTTTATGANGCTTGTATTCCTGCTGGAT 1020
 QY 1021 TTCTTTGAGCAATGTCAGAGAAATTTGATTGGCTCAGGTTTATGAGCTTCTGCTGATT 1080
 DB 1021 TTCTTTGAGCAATGTCAGAGAAATTTGATTGGCTCAGGTTTATGAGCTTCTGCTGATT 1080
 QY 1081 ATCTTCATGTTGCTGCTGTCGGAACATTCATGCTTAAAGCAACGCTGCTCCTC 1140
 DB 1081 ATCTTCATGTTGCTGCTGTCGGAACATTCATGCTTAAAGCAACGCTGCTCCTC 1140
 QY 1141 TTGTTATCTCGTTTATGGGTGCTGCTATTTGTTGCTACAGCCCTTGAATATAT 1200
 DB 1141 TTGTTATCTCGTTTATGGGTGCTGCTATTTGTTGCTACAGCCCTTGAATATAT 1200
 QY 1201 TCTCTTTTACTATAGAAATGCTGAGTGGGACAGCTGAAATATATATAA 1254
 DB 1201 TCTCTTTTACTATAGAAATGCTGAGTGGGACAGCTGAAATATATATAA 1254

RESULT 3

US-10-011-771B-1
 ; Sequence 1, Application US/10011771B
 ; Publication No. US20030082780A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Weigel, Paul H.
 ; APPLICANT: DeAngelis, Paul
 ; APPLICANT: Kumari, Kshama
 ; TITLE OF INVENTION: Hyaluronan Synthase Gene and Uses Thereof
 ; FILE REFERENCE: 3554.011
 ; CURRENT APPLICATION NUMBER: US/10/011,771B
 ; CURRENT FILING DATE: 2001-10-11
 ; PRIOR APPLICATION NUMBER: US 09/178,851
 ; PRIOR FILING DATE: 1998-10-26
 ; PRIOR APPLICATION NUMBER: US 60/064,435
 ; PRIOR FILING DATE: 1997-10-31
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 1254
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 US-10-011-771B-1

Query Match 100.0% Score 1254; DB 14; Length 1254;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGAGACATATATAAACCCTCAATACCTGTTGGCTTTAGTATTTTGGGTACTGTG 60
 DB 1 ATGAGACATATATAAACCCTCAATACCTGTTGGCTTTAGTATTTTGGGTACTGTG 60

QY 61 ATTACGTCAATGTTATCTCTTTGGTCTAAAGGAAGCTTGTCAATTTATGCTTTTG 120
 DB 61 ATTACGTCAATGTTATCTCTTTGGTCTAAAGGAAGCTTGTCAATTTATGCTTTTG 120
 QY 121 CTGATAGCTTACCTTATAGTCAAAATGTCCTTATCCCTTTTACAAGCCATTTAAGGA 180
 DB 121 CTGATAGCTTACCTTATAGTCAAAATGTCCTTATCCCTTTTACAAGCCATTTAAGGA 180
 QY 181 AGGCTTGGGCAATATAAGTGTGGAGCCATTAATCCCTCTTATTAAGAGAGATGCTGATCA 240
 DB 181 AGGCTTGGGCAATATAAGTGTGGAGCCATTAATCCCTCTTATTAAGAGAGATGCTGATCA 240
 QY 241 TTCTTAGAGAGCTTAAAGAGTGTTCAGCAGCAAACTATCCCTTAGCAGAAATTTATGTT 300
 DB 241 TTCTTAGAGAGCTTAAAGAGTGTTCAGCAGCAAACTATCCCTTAGCAGAAATTTATGTT 300
 QY 301 GTTGACGATGGAAGTGTGATGAGAGAGTATTAAGCGCATTTAAGAGACTATGCTGAGAC 360
 DB 301 GTTGACGATGGAAGTGTGATGAGAGAGTATTAAGCGCATTTAAGAGACTATGCTGAGAC 360
 QY 361 ACTGGTGACCTATCAAGCAATGTCATGTTTCATCGTCAGAGAAAATCAAGAAAGCGT 420
 DB 361 ACTGGTGACCTATCAAGCAATGTCATGTTTCATCGTCAGAGAAAATCAAGAAAGCGT 420
 QY 421 CATGCAAGGCTGGGCTTTGAAAGATCAGACGCTGATGCTCTTTTTCACGCTTGACTCA 480
 DB 421 CATGCAAGGCTGGGCTTTGAAAGATCAGACGCTGATGCTCTTTTTCACGCTTGACTCA 480
 QY 481 GATACTTATATCTACCTCTGATGCTTTTAGAGAGTGTGTTTAAACCTTTAATGAGCCAACT 540
 DB 481 GATACTTATATCTACCTCTGATGCTTTTAGAGAGTGTGTTTAAACCTTTAATGAGCCAACT 540
 QY 541 GTTTTTCGTCGCAAGGCTCACCCTTAATGTCAGAAATAGACAAACCAATCTCTTAACACG 600
 DB 541 GTTTTTCGTCGCAAGGCTCACCCTTAATGTCAGAAATAGACAAACCAATCTCTTAACACG 600
 QY 601 TTGACAGATATTCGCTATGATAATGCTTTTGGCGTTGAACGAGCTGCCCAATCCGTTACA 660
 DB 601 TTGACAGATATTCGCTATGATAATGCTTTTGGCGTTGAACGAGCTGCCCAATCCGTTACA 660
 QY 661 GGTAAATATCTTTGCTCAGGTCGCTTAGCTTTACAGACGCGAGGTGTTTCCT 720
 DB 661 GGTAAATATCTTTGCTCAGGTCGCTTAGCTTTACAGACGCGAGGTGTTTCCT 720
 QY 721 AACATAGATAGATACATCAACAGACGCTTCCCTGGGTATTCCTGTAAGTATGGTGATGAC 780
 DB 721 AACATAGATAGATACATCAACAGACGCTTCCCTGGGTATTCCTGTAAGTATGGTGATGAC 780
 QY 781 AGGTGCTTGACCACTATGCAACTGATTAGGAAGAGCTGTTTATCAATCCACTGCTAAA 840
 DB 781 AGGTGCTTGACCACTATGCAACTGATTAGGAAGAGCTGTTTATCAATCCACTGCTAAA 840
 QY 841 TGTAATACAGATGTTCCCTGACAAGATGCTTACTTCTGTAAGAGAGAGCTGTTTATCAATCCACTGCTAAA 900
 DB 841 TGTAATACAGATGTTCCCTGACAAGATGCTTACTTCTGTAAGAGAGAGCTGTTTATCAATCCACTGCTAAA 900
 QY 901 AAGTCCCTTTTAGAGAGTCCATTAATTCGTTAGGAAGATCATGAACAAATCCTTTGTA 960
 DB 901 AAGTCCCTTTTAGAGAGTCCATTAATTCGTTAGGAAGATCATGAACAAATCCTTTGTA 960
 QY 961 GGCCTATGAGCACTACTGAGGTGCTATGTTTATGANGCTTGTATTCCTGCTGGAT 1020
 DB 961 GGCCTATGAGCACTACTGAGGTGCTATGTTTATGANGCTTGTATTCCTGCTGGAT 1020
 QY 1021 TTCTTTGAGCAATGTCAGAGAAATTTGATTGGCTCAGGTTTATGAGCTTCTGCTGATT 1080
 DB 1021 TTCTTTGAGCAATGTCAGAGAAATTTGATTGGCTCAGGTTTATGAGCTTCTGCTGATT 1080
 QY 1081 ATCTTCATGTTGCTGCTGTCGGAACATTCATGCTTAAAGCAACGCTGCTCCTC 1140
 DB 1081 ATCTTCATGTTGCTGCTGTCGGAACATTCATGCTTAAAGCAACGCTGCTCCTC 1140

QY 1141 TTGTTATCTCCGTTTATGAGGCTGCTGCTATTTGTTTCTGACAGCCCTTGAATATAT 1200
 Db 1141 TTGTTATCTCCGTTTATGAGGCTGCTGCTATTTGTTTCTGACAGCCCTTGAATATAT 1200
 QY 1201 TCTCTTTTACTATTAGAAATGCTGCTGAGGGAACACGTAATAAATATATAA 1254
 Db 1201 TCTCTTTTACTATTAGAAATGCTGCTGAGGGAACACGTAATAAATATATAA 1254

RESULT 4

US-10-172-527-1
 ; Sequence 1, Application US/10172527
 ; Publication No. US2003092118A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Weigel, Paul H
 ; APPLICANT: Kumari, Kshama
 ; APPLICANT: DeAngelis, Paul
 ; TITLE OF INVENTION: HYALURONAN SYNTHASE GENES AND EXPRESSION THEREOF IN BACILLUS SUBT
 ; FILE REFERENCE: 3554.048
 ; CURRENT APPLICATION NUMBER: US/10/172,527
 ; CURRENT FILING DATE: 2002-06-13
 ; PRIOR APPLICATION NUMBER: 60/297,788
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: 60/297,744
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: 09/469,200
 ; PRIOR FILING DATE: 1999-12-21
 ; PRIOR APPLICATION NUMBER: 09/178,851
 ; PRIOR FILING DATE: 1998-10-26
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 1254
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 US-10-172-527-1

Query Match 100.0%; Score 1254; DB 14; Length 1254;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGAGAACATTAATAAACCTCATACCTGTTGGCTTTAGTATTTTGGGTACTGTTG 60
 Db 1 ATGAGAACATTAATAAACCTCATACCTGTTGGCTTTAGTATTTTGGGTACTGTTG 60
 QY 61 ATTACGTCATGTTTATCTCTTTGGTCTAAAGAGCTGTCAATTTATGCTTTTGG 120
 Db 61 ATTACGTCATGTTTATCTCTTTGGTCTAAAGAGCTGTCAATTTATGCTTTTGG 120
 QY 121 CTGATAGCTTACCTATTAGTCAAAATGTCCTTATCCCTTTTACAGGCAATTAAGGGA 180
 Db 121 CTGATAGCTTACCTATTAGTCAAAATGTCCTTATCCCTTTTACAGGCAATTAAGGGA 180
 QY 181 AGGGCTGGCAATATAGTCTGACCAATATTCCTCTTATACGAAGATGCTGATCA 240
 Db 181 AGGGCTGGCAATATAGTCTGACCAATATTCCTCTTATACGAAGATGCTGATCA 240
 QY 241 TTGCTAGAGACTTAAAAAGTGTTCAGCAGCAAACTATCCCTTAGCAGAAATTTATGTT 300
 Db 241 TTGCTAGAGACTTAAAAAGTGTTCAGCAGCAAACTATCCCTTAGCAGAAATTTATGTT 300
 QY 301 GTTGACGATGGAAGTGTGATGAGACAGGTATTAAAGCGATTGAAGACTATGCGGTGAC 360
 Db 301 GTTGACGATGGAAGTGTGATGAGACAGGTATTAAAGCGATTGAAGACTATGCGGTGAC 360
 QY 361 ACTGGTGACCTATCAAGCAATGTCATTCCTGCTGACAGAAATCAAGAAAGCT 420
 Db 361 ACTGGTGACCTATCAAGCAATGTCATTCCTGCTGACAGAAATCAAGAAAGCT 420
 QY 421 CATGCACAGGCTGGGCTTTGAAAGATCAGACGCTGATGCTTTTGGACCTGACTCA 480
 Db 421 CATGCACAGGCTGGGCTTTGAAAGATCAGACGCTGATGCTTTTGGACCTGACTCA 480

QY 481 GATACCTATCTACCTGATGCTTTAGAGAGTGTGTTAAAAAACCCTTTAATGACCAACT 540
 Db 481 GATACCTATCTACCTGATGCTTTAGAGAGTGTGTTAAAAAACCCTTTAATGACCAACT 540
 QY 541 GTTTTGTGCGAGGGTCACTTAAATCTCAGAAATAGACAAACCAATCTCTTTAACAGC 600
 Db 541 GTTTTGTGCGAGGGTCACTTAAATCTCAGAAATAGACAAACCAATCTCTTTAACAGC 600
 QY 601 TTGACAGATATTCGCTATGATATGCTTTTGGCGTTGAACGAGCTGCCAATCCGTTACA 660
 Db 601 TTGACAGATATTCGCTATGATATGCTTTTGGCGTTGAACGAGCTGCCAATCCGTTACA 660
 QY 661 GGTATATCTCTTGTGCTCAGGTCCGCTTACGCTTTACAGACGCGAGGTGTTGCTCT 720
 Db 661 GGTATATCTCTTGTGCTCAGGTCCGCTTACGCTTTACAGACGCGAGGTGTTGCTCT 720
 QY 721 AACATAGATAGATACATCAACAGACCTTCTGGGTATTCCTGTAAGTATGTTGATGAC 780
 Db 721 AACATAGATAGATACATCAACAGACCTTCTGGGTATTCCTGTAAGTATGTTGATGAC 780
 QY 781 AGGTGCTTGACCAACTATGCACTGATTTAGGAAAGACTGTTTATCAATCCACTGCTAAA 840
 Db 781 AGGTGCTTGACCAACTATGCACTGATTTAGGAAAGACTGTTTATCAATCCACTGCTAAA 840
 QY 841 TGTATTACAGATGTTTCTCTGACAGATGCTTACTTACTTGAAGCAGCAAAACCGCTGGAAC 900
 Db 841 TGTATTACAGATGTTTCTCTGACAGATGCTTACTTACTTGAAGCAGCAAAACCGCTGGAAC 900
 QY 901 AAGTCTCTTCTTAGAGAGTCCATTTCTGTTAAGAAATCATGAACAATCCCTTTGTA 960
 Db 901 AAGTCTCTTCTTAGAGAGTCCATTTCTGTTAAGAAATCATGAACAATCCCTTTGTA 960
 QY 961 GCCTTATGACCACTTGTAGGTGCTGATGTTTATGCTGTTTATCTGTTGGTGGAT 1020
 Db 961 GCCTTATGACCACTTGTAGGTGCTGATGTTTATGCTGTTTATCTGTTGGTGGAT 1020
 QY 1021 TTCTTTTAGGCAATGTCAGAGAAATTTGATGGCTCAGGGTTTTAGCCTTTCTGTTGAT 1080
 Db 1021 TTCTTTTAGGCAATGTCAGAGAAATTTGATGGCTCAGGGTTTTAGCCTTTCTGTTGAT 1080
 QY 1081 ATCTTATGTTGCTGCTGCGACATTCATTTACATGCTTAAAGCAGCCGCTGCTCTTC 1140
 Db 1081 ATCTTATGTTGCTGCTGCGACATTCATTTACATGCTTAAAGCAGCCGCTGCTCTTC 1140
 QY 1141 TTGTTATCTCCCTTTTATGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 Db 1141 TTGTTATCTCCCTTTTATGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 QY 1201 TCTCTTTTACTATTAGAAATGCTGACGCGGGAACACGTAATAAATTTATATA 1254
 Db 1201 TCTCTTTTACTATTAGAAATGCTGACGCGGGAACACGTAATAAATTTATATA 1254

RESULT 5

US-10-326-185-1
 ; Sequence 1, Application US/10326185
 ; Publication No. US20030175902A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sloma, Alan
 ; APPLICANT: Behr, Regine
 ; APPLICANT: Widner, William
 ; APPLICANT: Tang, Maria
 ; APPLICANT: Sternberg, David
 ; APPLICANT: Brown, Stephen
 ; TITLE OF INVENTION: Methods for Producing Hyaluronan In a Recombinant Host Cell
 ; FILE REFERENCE: 10241.200-US
 ; CURRENT APPLICATION NUMBER: US/10/326,185
 ; CURRENT FILING DATE: 2002-12-20
 ; PRIOR APPLICATION NUMBER: US 60/342,644
 ; PRIOR FILING DATE: 2001-12-21
 ; NUMBER OF SEQ ID NOS: 108
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1

```

; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1251)
; OTHER INFORMATION:
US-10-326-185-1

Query Match      97.7%; Score 1225.4; DB 12; Length 1251;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1235; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 ATGAGACATTAATAAACCTCATACCTGTTGGCCCTTTAGTATTTTGGGTACGTTG 60
Db 1 ATGAGACATTAATAAACCTCATACCTGTTGGCCCTTTAGTATTTTGGGTACGTTG 60

QY 61 ATTACGCTCAATGTTATCTCTTTGGTCTAAAGGAAGCTTGCTCAATTTATGCTTTG 120
Db 61 ATTACGCTCAATGTTATCTCTTTGGTCTAAAGGAAGCTTGCTCAATTTATGCTTTG 120

QY 121 CTGATAGCTTACCTATTAGTCAAAATGCTTATCTTTTAAAGCAAGCTTGCTCAATTTATGCTTTG 180
Db 121 CTGATAGCTTACCTATTAGTCAAAATGCTTATCTTTTAAAGCAAGCTTGCTCAATTTATGCTTTG 180

QY 181 AGGGCTGGCAATATAAGGTTGCGCCATTATCCCTCTTATAAGGAAGATGCTGATCA 240
Db 181 AGGGCTGGCAATATAAGGTTGCGCCATTATCCCTCTTATAAGGAAGATGCTGATCA 240

QY 241 TTGCTAGAGACCTTAAAGTGTTCAGCAGCAAACTATCCCTTAGCAGAAATTTATGTT 300
Db 241 TTGCTAGAGACCTTAAAGTGTTCAGCAGCAAACTATCCCTTAGCAGAAATTTATGTT 300

QY 301 GTTGACGATGAAGTGTCTGATGAGCAGAGTATTAAGCGCATTTAGGACTATGCGGTGAC 360
Db 301 GTTGACGATGAAGTGTCTGATGAGCAGAGTATTAAGCGCATTTAGGACTATGCGGTGAC 360

QY 361 ACTGCTGACCTTCAAGCAATGCTATGTTTCATCGGTGAGAGAAATCAAGGAAGGTT 420
Db 361 ACTGCTGACCTTCAAGCAATGCTATGTTTCATCGGTGAGAGAAATCAAGGAAGGTT 420

QY 421 CATGACAGGCTGGGCTTTCAAGATGATGAGCTGATGCTTTTGGCCCTTACCTGATCA 480
Db 421 CATGACAGGCTGGGCTTTCAAGATGATGAGCTGATGCTTTTGGCCCTTACCTGATCA 480

QY 481 GATCTATATCTACCTGATGCTGTTAGAGAGGTTGTTAAACCTTTAATGACCCACT 540
Db 481 GATCTATATCTACCTGATGCTTAAAGAGGTTGTTAAACCTTTAATGACCCACT 540

QY 541 GTTTTCTCGGAGGGTCACTTAATGTCAGAAATAGACAAACCAATCTTTAACAGC 600
Db 541 GTTTTCTCGGAGGGTCACTTTAATGTCAGAAATAGACAAACCAATCTTTAACAGC 600

QY 601 TTGACAGATATTCGCTATGATAATGCTTTTGGCTTGAAGCAGCTGCCCAATCCGTACA 660
Db 601 TTGACAGATATTCGCTATGATAATGCTTTTGGCTTGAAGCAGCTGCCCAATCCGTACA 660

QY 661 GGTAAATCCCTGTTGCTCAGGTCGCTTACGCTTACAGCGGAGGTGCTGCTTCT 720
Db 661 GGTAAATCCCTGTTGCTCAGGTCGCTTACGCTTACAGCGGAGGTGCTGCTTCT 720

QY 721 AACATAGATAGATACATCAACGACCTTCTTGGGTATTCCTGTAAGTATTTGATGAC 780
Db 721 AACATAGATAGATACATCAACGACCTTCTTGGGTATTCCTGTAAGTATTTGATGAC 780

QY 781 AGGTGCTTGACCAATATGCACTGATTTAGGAAGAGCTTTTATCAATCCACTGCTAAA 840
Db 781 AGGTGCTTGACCAATATGCACTGATTTAGGAAGAGCTTTTATCAATCCACTGCTAAA 840

QY 841 TGTATTACAGATGTTCCCTGACAGATGCTACTTTACTTTGAAGCAGCAAAACCGTGGAC 900
Db 841 TGTATTACAGATGTTCCCTGACAGATGCTACTTTACTTTGAAGCAGCAAAACCGTGGAC 900

```

```

QY 901 AAGTCTCTCTTTAGAGAGTCCATTAATTTCTGTTAAGAAATCATGAACAATCCCTTTGTA 960
Db 901 AAGTCTCTCTTTAGAGAGTCCATTAATTTCTGTTAAGAAATCATGAACAATCCCTTTGTA 960

QY 961 GCCCTATGGACCATATCTTGAAGTGTCTATGTTTATGATGCTGTTTATTCGTGTTGAT 1020
Db 961 GCCCTATGGACCATATCTTGAAGTGTCTATGTTTATGATGCTGTTTATTCGTGTTGAT 1020

QY 1021 TTTCTTTAGGCAATGTCAGAGAAATTTGATGCTCAGGCTTTTACGCTTTCTGTTGAT 1080
Db 1021 TTTCTTTAGGCAATGTCAGAGAAATTTGATGCTCAGGCTTTTACGCTTTCTGTTGAT 1080

QY 1081 ATCTTCATTTGCTCTTTGTCGTAATATTTCACTATATGCTTAAGCACCCTGCTCTTC 1140
Db 1081 ATCTTCATTTGCTCTTTGTCGTAATATTTCACTATATGCTTAAGCACCCTGCTCTTC 1140

QY 1141 TTGTTATCCCTTTTATGCGGTGCTGCAATTTGTTTCTCAGCCCTTGAATATAT 1200
Db 1141 TTGTTATCCCTTTTATGCGGTGCTGCAATTTGTTTCTCAGCCCTTGAATATAT 1200

QY 1201 TCTCTTTTACTATTAGAAATGCTGACTGGGGAACAGCTAAATAATTA 1251
Db 1201 TCTCTTTTACTATTAGAAATGCTGACTGGGGAACAGCTAAATAATTA 1251

```

RESULT 6

```

US-10-326-185-108
; Sequence 108, Application US/10326185
; Publication No. US20030175902A1
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; APPLICANT: Behr, Regine
; APPLICANT: Widner, William
; APPLICANT: Tang, Maria
; APPLICANT: Sternberg, David
; APPLICANT: Brown, Stephen
; TITLE OF INVENTION: Methods for Producing Hyaluronan In a Recombinant Host Cell
; FILE REFERENCE: 10241.200-US
; CURRENT APPLICATION NUMBER: US/10/326,185
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/342,644
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 108
; LENGTH: 5158
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-10-326-185-108

```

```

Query Match      91.9%; Score 1152; DB 12; Length 5158;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 TCAATTAATGCTTTTGTGCTAGTACCTATAGTCAAAATGCTCTTATCCCTTTT 162
Db 1 TCAATTAATGCTTTTGTGCTAGTACCTATAGTCAAAATGCTCTTATCCCTTTT 60

QY 163 TACAGCCATTTAAGGAAGGCTGGGCAATATAAGTTGCGAGCCATTTATCCCTTAT 222
Db 61 TACAGCCATTTAAGGAAGGCTGGGCAATATAAGTTGCGAGCCATTTATCCCTTAT 120

QY 223 AACGAGATGCTGAGTCATGCTAGACCTTAAAGTGTTCAGCAGCAAACTATCC 282
Db 121 AACGAGATGCTGAGTCATGCTAGACCTTAAAGTGTTCAGCAGCAAACTATCC 180

QY 283 CTAGCAGAAATTTATGTTGACGATGGAAGTCTGATGAGCAGGATTAAGCGCAT 342
Db 181 CTAGCAGAAATTTATGTTGACGATGGAAGTCTGATGAGCAGGATTAAGCGCAT 240

QY 343 GAAGACTATGCGTGACACTGCTGACCTATCAAGCAATGCTATTGTCATCGGCAG 402
Db 241 GAAGACTATGCGTGACACTGCTGACCTATCAAGCAATGCTATTGTCATCGGCAG 300

```

```

QY 403 AAAAAATCAAGAAAGCCTCATGACAGAGCCTGGGCGCTTTGAAAGATCAGACGCTGATGTC 462
Db 301 AAAAAATCAAGAAAGCCTCATGACAGAGCCTGGGCGCTTTGAAAGATCAGACGCTGATGTC 360
QY 463 TTTTTCAGCGTGTGACTCAGATACATATATCTACCTGATGCTTTAGAGGAGTGTGTTAAAA 522
Db 361 TTTTTCAGCGTGTGACTCAGATACATATATCTACCTGATGCTTTAGAGGAGTGTGTTAAAA 420
QY 523 ACCTTTAATGACCCAACTGTTTTCGTCGACGGGTCACCTTATGTCAGAAATAGACAA 582
Db 421 ACCTTTAATGACCCAACTGTTTTCGTCGACGGGTCACCTTATGTCAGAAATAGACAA 480
QY 583 ACCAATCTCTTAACAGCGTTGACAGATATTCGCTATGATATGCTTTTGGCGTTGAACGA 642
Db 481 ACCAATCTCTTAACAGCGTTGACAGATATTCGCTATGATATGCTTTTGGCGTTGAACGA 540
QY 643 GTCGCCCAATCCGTTACAGTAATATCTTGTGTCGACGGGTCAGCTCGCTTACAGGTTACAGA 702
Db 541 GTCGCCCAATCCGTTACAGTAATATCTTGTGTCGACGGGTCAGCTCGCTTACAGGTTACAGA 600
QY 703 CCGGAGTGTGTTTCTTAAATAGATAGATACATCAACAGACCTTCTTGGGTATTCCT 762
Db 601 CCGGAGTGTGTTTCTTAAATAGATAGATACATCAACAGACCTTCTTGGGTATTCCT 660
QY 763 GTAAATATGCTGATGACAGGTGCTTGACCACTATGCAACTGATTTAGGAAAGACTGTT 822
Db 661 GTAAATATGCTGATGACAGGTGCTTGACCACTATGCAACTGATTTAGGAAAGACTGTT 720
QY 823 TATCAATCCACTGCTAAATGATTAACAGATGTTCTGCAAGATGCTACTACTTGAAG 882
Db 721 TATCAATCCACTGCTAAATGATTAACAGATGTTCTGCAAGATGCTACTACTTGAAG 780
QY 883 CAGCAAAACCGCTGGAAAGTCTTCTTTAGAGAGTCCATTTATTCGTTAAGAAATC 942
Db 781 CAGCAAAACCGCTGGAAAGTCTTCTTTAGAGAGTCCATTTATTCGTTAAGAAATC 840
QY 943 ATGCAACATCTTTTCTAGCCCTATGACCATATCTAGGTCGCTATGTTTATGAGCTT 1002
Db 841 ATGCAACATCTTTTCTAGCCCTATGACCATATCTAGGTCGCTATGTTTATGAGCTT 900
QY 1003 GTTATATCTGCTGGTGAATTTCTTTTAGGCAATGTCAGAAATTTGATGGCTCAGGCTT 1062
Db 901 GTTATATCTGCTGGTGAATTTCTTTTAGGCAATGTCAGAAATTTGATGGCTCAGGCTT 960
QY 1063 TTAGCCCTTTCTGCTGATATCTTCAATGTTGCCCTGTGTCGGAACATTCATTCATGCTT 1122
Db 961 TTAGCCCTTTCTGCTGATATCTTCAATGTTGCCCTGTGTCGGAACATTCATTCATGCTT 1020
QY 1123 AAGCACCGCTGCTCTTCTTCTTATCTCCGCTTTTATGGGTCGCTGATTTGTTGCTTA 1182
Db 1021 AAGCACCGCTGCTCTTCTTCTTATCTCCGCTTTTATGGGTCGCTGATTTGTTGCTTA 1080
QY 1183 CAGCCCTTGAATATATATCTTTTACTATAGAAATGCTGACTGGGGAACACGTTAA 1242
Db 1081 CAGCCCTTGAATATATATCTTTTACTATAGAAATGCTGACTGGGGAACACGTTAA 1140
QY 1243 AAATATATATAA 1254
Db 1141 AAATATATATAA 1152

```

RESULT 7

```

us-10-326-185-102
; Sequence 102, Application US/10326185
; Publication No. US20030175902A1
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; APPLICANT: Behr, Regine
; APPLICANT: Widner, William
; APPLICANT: Tang, Maria
; APPLICANT: Sternberg, David
; APPLICANT: Brown, Stephen

```

```

; TITLE OF INVENTION: Methods for Producing Hyaluronan In a Recombinant Host Cell
; FILE REFERENCE: 10241.200-US
; CURRENT APPLICATION NUMBER: US/10/326,185
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/342,644
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 102
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Streptococcus uberis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1251)
; OTHER INFORMATION:
; US-10-326-185-102

Query Match 49.7%; Score 623.2; DB 12; Length 1251;
Best Local Similarity 69.4%; Pred. No. 3.9e-167;
Matches 863; Conservative 0; Mismatches 378; Indels 3; Gaps 1;

QY 1 ATGAGAACATTAATAAACCTCATACCTGTGTGGCGCTTTAGTATTTTGGGTACTGTTG 60
Db 1 ATGAGAACATTAATAAACCTCATACCTGTGTGGCGCTTTAGTATTTTGGGTACTGTTG 60
QY 61 ATTAGCTCAATGTTATCTCTTTGGTGTCTAAAGGAAGCTGTCTCAATTTATGCTTTT 120
Db 61 ATTGGGCTTAATGTTTGTATTTGGAATTAAGGAAGTCTAACAGTGTATGGGATTT 120
QY 121 CTGATAGCTTACCTATTAGTCAAAATGCTTATCCCTTTTACAGCCATTTAAGGA 180
Db 121 CTATTAACCTATTGTCGATAAATGGGATTAATCTTTTTTTATCGCCCTATAAGGA 180
QY 181 AGGCTGGGCAATTAAGSTTGCAGCCATTATCCCTCTTATAAGGAAGTGTGAGTCA 240
Db 181 AGTGTAGTCAATTAAGSTTGCAGCTATTATCCATCTTATAAGGAGTGTGCTG 240
QY 241 TTCTAGAGACTTAAAGAGTCTTCCAGCAGCAACCTATCCCTTAGCAGAAATTTATGTT 300
Db 241 TTACTAGAACTTAAAGAGTGTCTAAAGCAACATATCCAAATTCGAGAAATTTTCGTA 300
QY 301 GTTGACATGGAAGTGTCTGATGAGACAGGTATTAAAGCATTTGAAGACTATGTGCTGAC 360
Db 301 ATTGACATGAGTGTCTGATGATTAAGACAGGTATTAAATGTGTCGAGACTATGTGAAGTTA 360
QY 361 ACTGTGACCTTATCAGCAATGTCTTCTTATCGGTGTCAGAGAAATCAAGGAAGGCT 420
Db 361 AATGG---CTTTGGAGACCAAGTTATCGTTCATCAGATCGCTGAAATTTTGGTAAAGA 417
QY 421 CATGCACAGGCTGGGCGCTTTGAAAGATCAGAGCTGATGCTTTTGGACGCTGACTCA 480
Db 418 CATGCTCAGGCTTGGGCAATTTGAAGGTCGTGATGCTTTTCTTAAACAGTGGATTCA 477
QY 481 GATACTTATCTACCCCTGATGCTTTTAGAGAGTGTCTTAAAGAACTTTAATGACCCAACT 540
Db 478 GATACCTACATCTATCTGCTGCTCTTGAAGAATTTAATAGACATTTAATGATCCAGAG 537
QY 541 GTTTTTGTCGAGGCTGACCTTAAGTCAAGAAATAGACAAACCAATCTCTTAACAGCG 600
Db 538 GTCTACGCTGCACTGCTCATTTAAATGCAAGAAATAGACAAACCAATCTCTTAACAG 597
QY 601 TTGACAGATATTCGCTATGATTAATGCTTTTGGGTTGAGCAGCTGCCCAATCCGTTACA 660
Db 598 CTGACTGATATTCGTTACGATTAAGCAATTTGGTGTAGAACGCTGCTGCTGCTGTACG 657
QY 661 GGTAAATCTCTTGTGTTTCTCAGGTCCTTGTAGCGTTTACAGACGCGAGGTGTTCTTCT 720
Db 658 GGAATAATTTTGGTTTGTCCGGACCTTTAAGTATTTATAGAGCTTCCTGCGGTATGCA 717
QY 721 AACATAGATAGATACATCAACAGACCTTCCCTGGGTATTCCTGTAAGTATTTGGTATGAC 780
Db 718 AATCTTGACGCTATACCTCAACAACTTTCTTGTGCTCCCTGTGAAGCATAGGGATGAC 777

```


QY	10	TTAAAAACCTCATAACTGTTGGCTTGTAGTATTTTTTGGGTACTGTGGATTAACGTC	69
Db	202	TTTAAAAAAACCTTTAAATTTGTTTATCCTCTTTATTTTTTTTGATATCTATCTTGATTTATCTA	261
QY	70	AATGTTTATCTCTTTGGTCTAAAGGAAGCTTGTCAATTTATGGCTTTTGGCTGTATAGCT	129
Db	262	AATATGTATCTATTTTGGAAC---ATCACTGTAGGAATTTATGGAGTAAATTAATTAATTAAC	318
QY	130	TACCTATTACTCAAAATGCTCTATCCTTTTTTACAAGCCATTTAAAGGAAGCGCTGGG	189
Db	319	TATCTAGTTATCAAACTTGGATATCTTTCCCTTTATGAGCCATTTAAAGGNAATCCAAAT	378
QY	190	CAATATAAGTGTGCGACCAATTATTCCTCTTATACGAAGATGCTGAGTCATGCTAGAG	249
Db	379	GACTATAAAGTGTGCTGCTGTAATTCCTCTTATAATGAAGATCGCGAGTCATTTAGAA	438
QY	250	ACCTTTAAAAAGTGTTCAGCAGCAAACTATCCCTACGACGAATTTATGTTGTTGACAT	309
Db	439	ACACTTAAGTGTGTAGCAGACGCTATCCGTTTCAGAAATTTATTTGTTGATGAT	498
QY	310	GGAGTGTCTGATGAGACAGSTATTAAAGCGCATTAAGACATATGCTGCTGCACATCGTGAC	369
Db	499	GGGAGTTCAACACACAGATGCAATAACAATTAATTGAAGAGTATGTAATTAAGNAGTGTGAT	558
QY	370	CTATCAGCAATGCTCATTTGTTTCATCGGTACAGAAATAATCAAGGAAGCGTCAATGCACAG	429
Db	559	ATTGTCGAAAGCTTATCGTTCACCTCCCTCTGCAATAAAGGAAACGCCATGCTCAA	618
QY	430	GCCTGGGCGCTTTGAAAGATCAGACGCTGAATGCTTTTTGACGTTGACTCAGATCACTTAT	489
Db	619	GGGTGGSCATTTGAAAGATCTGACGCTGACGTTTTTTTTAACGTAGACTCAGATCTTAT	678
QY	490	ATCTACCTCGATGCTTTTAGAGGAGTGTGTTTAAAAACCTTTAAATGACCCAACCTGTTTTGCT	549
Db	679	ATCTATCCAAATGCGCTTAGAAGAACTCCCTAAAAAGCTTCAATGATGAGACAGTTATGCT	738
QY	550	GGCAGGGTCACCTTAATGTCAAGAAATAGCAAAACCAATCTCTTAACACGCTTGACAGAT	609
Db	739	GCACAGGACATTTCAATGCTAGAAACAGACAACTAACTTATTAAACGGACATACAGAT	798
QY	610	ATTCGCTATGATAATGCTTTTGGCGTTGAAGAGCTGCCCAATCCGTTACAGGTAATATC	669
Db	799	ATCCGTTATACGATAATGCCCTTTGGGTGGAGCGTCTGCTCAATCATTAACAGGTAAATTT	858
QY	670	CYTGTGTTGCTCAGGTCGCGCTTACGTTTTACAGACGCGAGGTGGTGTGTTTCCCTAACATAGAT	729
Db	859	TTAGTTTGCTTCAGGACCAATGAGATATTATCGAGGTCAAGTGAATTTCTCACTACTAGAG	918
QY	730	AGATACATCAACGACAGCTTCCTTGGGTATTCCTTAAGTATTGGTGAATGACAGGTGCTTG	789
Db	919	CGCTATAAAAAATCAAAACATTCCTTAGTTTACCTGTTAGCAATTTGGGGATGATCGATGTTTA	978
QY	790	ACCAACTATGCAACTGATTTTAGAAGACTGTTTATCAATCCACTGCTAAATGTATTACA	849
Db	979	ACAAATTTATGCTATTGATTTAGACGCACTGCTTACCANTCAACAGCTAGTGTGTAFACT	1038
QY	850	GATGTCCTTGACAAGATCTACTTACTTTGAGGACGCAAAACCGCTGGAAAGTCCCTTC	909
Db	1039	GATGTACCTTTCCAATTTAAAAAGTATTTTAAAGCAACAAAATTCGATGGGAATAAACTCTTTT	1098
QY	910	TTTAGAGAGTCCATTTATTTCTGTTAAGAAAAATCATGAACAATCCCTTTTGTGACCCATAGG	969
Db	1099	TTTAGAGATCTATTATTTCTGTTAAAAAAAATCTTTCTTAATCCCATCGCTTGCCCTATGG	1158

Db 439 ACCTTAAAGTGTGTAGACAGACCTATCCGTTATACAGAAATTTATTTGTGATGAT 498
QY 310 GGAAGTGTGATGAGACAGGATTAAGCGCATTAAGAGCTATGTCGGTGACACTGGTAC 369
Db 499 GGGAGTCAAAACACAGATGCATCAATTAATTAAGAGCTATTAATTAAGAGTGGAT 558
QY 370 CTAATCAAGCATGTCATGTCATCGGTACAGAAAAATCAAGAAAGCTCATGACAG 429
Db 559 ATTGTGCAAGCATGTCATGTCATCGGTACAGAAAAATCAAGAAAGCTCATGAC 618
QY 430 GCGTGGCTTTGAAGATCAGACGCTGATGCTTTTACCGTGTGACTCAGATCTAT 489
Db 619 GCGTGGCATTTGAAGATCTGACGCTGACGTTTTTAAACGCTAGACTCAGATCTAT 678
QY 490 ATCTACCTGATGCTTTAGAGGATGTTTAAACCTTTTATGACCCCACTGTTTGTCT 549
Db 679 ATCTACCAATGCTTTAGAGGATGTTTAAACCTTTTATGACCCCACTGTTTGTCT 549
QY 550 GCGAGGCTCAGCTTAATGTCAGAAATAGACAAACCAATCTCTTAAACGCTTGACAGAT 609
Db 739 GCAACAGCATTTGAATCTAGAACAGACAACTATCTATTAAACGCTTGACAGAT 798
QY 610 ATTCGCTATGATGCTTTGCGGTTGAACGAGCTGCGCAATCGCTTACAGGTATATC 669
Db 799 ATCCGTTACGATATGCTTTGGGTTGGAGCTGCTGCTCAATCATTAACAGGTATAT 859
QY 670 CTTGTTTCTCAGTCCGCTTAGCGTTTACAGCGAGGTGTTTCTTCTTACATAGAT 729
Db 859 TTAGTTTCTCAGGACCATTTAGTATTTATCGAGCTGAAGTGATTTCTTAACTACAG 918
QY 730 AGATACATCAACAGACCTTCTGCGTATTCCTGTAAGTATGCTGATGACAGGTGCTTG 789
Db 919 CGCTATAAATCAACATCTCTAGTTTACCTGTTAGCATGGGATGATCGATGTTA 978
QY 790 ACCAATGATCACTGATTTAGAAAGACTGTTTATCAATCCACTGTAAGTATTAACA 849
Db 979 ACAATATGCTATGTTAGTAGGACCTGCTACCAATCAACAGCTAGTGTACT 1038
QY 850 GATGTTCTCAGAGATGCTACTTACTTGAACGAGCAAAACCGCTGGAACAGTCTTC 909
Db 1039 GATGATCTTCCATTAAGAGTATTTAAGCAACAAATCGATGGAATTAATCTTTT 1098
QY 910 TTTAGAGTCCATTTCTGTTAAGAAATCATGAACAAATCCTTTTGTAGCCCTAGG 969
Db 1099 TTTAGAGATCTATTATTTCTGTTAAGAAATCTTTCTATCCATCGTTCCTTATGG 1158
QY 970 ACCATCTGAGTGTCTATGTTATGAGCTGTTTATCTGTTGATTTCTTTGTA 1029
Db 1159 ACTATTTGCAAGTCTGTTATGTTATGATGTTGATGTCGCAATGGGAATCTTTGTTT 1218
QY 1030 GCAATGTCAGAAATTTGATGCTCAGGTTTATGAGCTGTTTATCTGTTGATTTCTTAT 1089
Db 1219 AATCAAGCTATTCATAGACCTTATTAACATTTTGTCTTTTATCCATCATCTTATC 1278
QY 1090 GTTGCCTGTGCGGAACATTCATATGATGTTTATGAGCAACCGCTGCTTCTTTGTTATCT 1149
Db 1279 GTTGCCTGTGCGGAACATTCATATGATGTTTATGAGCAACCGCTGCTTCTTTGTTATCT 1338
QY 1150 CCGTATGAGGCTGCTGATTTGTTGCTTACAGCCCTTGAATATATTTCTTTT 1209
Db 1339 CCTCTGATGAATATTAACCTTGTGTTGCTTACAGCCCTTGAATATATTTCTTTT 1398
QY 1210 ACTATTAGAATGCTGCTGAGGAAACAGTAAAAA 1244
Db 1399 ACCATTAAAAATACGATGAGGAAACAGTAAAAA 1433

RESULT 12

US-10-326-185-92

; Sequence 92, Application US/10326185

; Publication No. US20030175902A1

; GENERAL INFORMATION:

; APPLICANT: Sloma, Alan
; APPLICANT: Behr, Regine
; APPLICANT: Widner, William
; APPLICANT: Tang, Maria
; APPLICANT: Sternberg, David
; APPLICANT: Brown, Stephen
; FILE OF INVENTION: Methods for Producing Hyaluronan In a Recombinant Host Cell
; CURRENT APPLICATION NUMBER: US/10/326,185
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/342,644
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 92
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Streptococcus pyogenes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1257)
; OTHER INFORMATION:
US-10-326-185-92

Query Match 48.0%; Score 601.4; DB 12; Length 1257;

Best Local Similarity 68.5%; Pred. No. 6.6e-161;

Matches 846; Conservative 0; Mismatches 386; Indels 3; Gaps 1;

QY 10 TTAATAAACTCATACATGTTGGCTTTAGTATTTTGGGTACTGTTGATTTACGTC 69
Db 10 TTTAAAAAATCTTAATTTGTTTATCCCTTTATTTTGTATATCTATCTGATTTATCTA 69
QY 70 AATGTTTATCTCTTGTGCTTAAAGGAAGCTGTCATTTATTTGCTTTTCTGATAGCT 129
Db 70 AATATGATCTATTGGAAC---ATCACTGTAGGAATTTATGGAGTAATTAATAAACC 126
QY 130 TACCTATTAGTCAAAATGTCCTTTATCTTTTACAAGCCATTTAAGGAAGGCTGG 189
Db 127 TATCTAGTTTAACTTGGATATCTTCTTTATGAGCCATTTAAGGAAGGCTGG 186
QY 190 CAATATAAGTTGACGCAATATCCCTTTTATACGAGATGCTGAGTCAATGCTAGAG 249
Db 187 GATATAAGTTGCTGTTATCTTCTTTATTAATGAAGATGCGGATCATTTATAGAA 246
QY 250 ACCTTAAAGTGTTCAGCAAAACCTATCCCTTAGCAGAAATTTATGTTGTTGACAT 309
Db 247 ACTCTTAAAGTGTTCAGCAAAACCTATCCCTTAGCAGAAATTTATGTTGTTGAT 306
QY 310 GGAAGTGTGATGACAGAGTATTAAGCGCATTTAAGCAATGCTGCGTGACACTGGTAC 369
Db 307 GGAAGTGTCAACACAGATGCAATTAATTAAGAGTATGTAATAGAGAGTGGAT 366
QY 370 CTATCAAGCAATGTCATGTTTCATCGGTACAGAAAAATCAAGGAAGGCTGACAG 429
Db 367 ATTTGTCGAACGTTATCTGTTTACCGTTCCTGTCATTAAGGAAGAACCCCATGCTCA 426
QY 430 GCGTGGCTTTGAAGATCAGACGCTGATGCTTTTTGACCGTTGACTCAGATCTTAT 489
Db 427 GCGTGGCATTTGAAGATCTGACGCTGACGTTTTTTAACGCTAGATTCAGATCTAT 486
QY 490 ATCTACCTGATGCTTTAGAGAGTGTGTTAAAAACCTTTAATGACCCCACTGTTTGTCT 549
Db 487 ATCTATCAAAATGCTTTAGAGAACTCCTAAAAAGCTCAATGATGAGACAGTTTATGCT 546
QY 550 GCGAGGCTCACTTATGTCAGAAATAGACAAACCAATCTCTTAAACGCTTGACAGAT 609
Db 547 GCAACAGCATTTGATGCTAGAACAGCAAACTAATCTATTAAACGCTTGACAGAT 606
QY 610 ATTCGCTATGATGCTTTTGGCGTTGAACGAGCTGCGCAATCCCTTACAGCTTAATATC 669
Db 607 ATCCGTTACGATATGCTTTTGGGTTGGAGCTGCTGCTCAATCATTTACAGGTATAT 666
QY 670 CTTGTTTCTCAGTCCGCTTAGCGTTTACAGAGCGAGGTGTTGTTCTTCAACATAGAT 729

Db 667 TTAGTTTGGTCCAGGACCATTTAGTATTTATCGAGCTGAAGTGAATATTCCTTAACCTTAGAG 726
QY 730 AGATACATCAACAGACCTTCCTGGGTATTCCTGTAGTATTTGGTATGACAGGTGCTG 789
Db 727 CGGTATATAAATAACCAATCCCTAGGTTTACCTGTAGCATTTGGGATGATCGATGTTTA 786
QY 790 ACCAACTATGCAACTGATTTAGGAAGACTGTTTATCAATCCACCTGCTAAATGTTATACA 849
Db 787 ACARATATGCTATTTAGGACGCACTGTACCAATCAACAGCTAGATGATGATACT 846
QY 850 GATGTTCCGCAAGATGCTACTTACTTGAAGCAGCAAAACCGCTGGAACAGTCCCTTC 909
Db 847 GATGTAACCTTCCCAATTAAGTATTTAAAGCAACAAAATCGATGGAATAAATCTTT 906
QY 910 TTTAGAGATCCCATTTCTGTTAGAAATCATGAACAAATCCCTTTTGGAGCCCTATGG 969
Db 907 TTTAAGAAATCATTTATTTCTGTTAAATAATCTTTCTAATCCCATCGTTGGCTTATGG 966
QY 970 ACCATCTTGAAGGTCTATGTTTATGATGCTTGTATTTCTGTTGGATTTCTTTGTA 1029
Db 967 ACTATTTTGGAGTCTTATGTTTATGATGTTGATGTTGTCGAATTTGGGAATCTTTGTT 1026
QY 1030 GGCATGTCAGAAATTTGATTTGGCTCAGGTTTTTAAAGCTTTCTGCTGATTAATCTCAT 1089
Db 1027 AATCAAGCTATTCAATTAGACCTTTATTAACCTTTTGGCTTTTATCCATCATCTTTATC 1086
QY 1090 GTTGCCTCTGTCGGAACATTCATTAACATGTTTAAAGCCTTCTGCTGCTTTGTTATCT 1149
Db 1087 GTTGCTTATGCTGAATGTTTATATATGATCAATCAACATCTGCTGCTGTTTGTATCT 1146
QY 1150 CCGTTTATGSGGTGCTGCATTTGTTGCTCCTCAGCCCTTGAAATTAATATCTCTTTT 1209
Db 1147 CCTCTGATGAATATTACACTGTTTGTCTTACAGCCCTAAACCTTTATCTTTATGC 1206
QY 1210 ACTATAGAAATGCTGACTGGGACACAGTAAAAA 1244
Db 1207 ACCATTAAAAATACGAATGGGGAACACGTAAAAA 1241

RESULT 13
US-10-172-527-17
; Sequence 17, Application US/10172527
; Publication No. US20030092118A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H
; APPLICANT: Kumari, Kshama
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: HYALURONAN SYNTHASE GENES AND EXPRESSION THEREOF IN BACILLUS SUBT
; FILE REFERENCE: 3554.048
; CURRENT APPLICATION NUMBER: US/10/172,527
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/297,788
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 60/297,744
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 09/469,200
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Bacillus anthracis px01
US-10-172-527-17
Query Match 22.6%; Score 282.8; DB 14; Length 1200;
Best Local Similarity 57.8%; Pred. No. 5.9e-70;
Matches 503; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 383 TCATTGTTTCATCGTCAGAGAAAATCAAGGAAGCTCATCAGAGGCTGGGCTTTG 442

Db 255 TAATCGTACATCGCTTTACCTAAGAAATTCGCGGAAAGACATGCTCAATATATGGGCTTTA 314
QY 443 AAAGATCAGAGCGCTGATGTCCTTTTGTACCGCTTGACTCAGATACCTTATATATACCTGATG 502
Db 315 AAGCGAACACAGCAGATGCTATTGTTACCATTTGATTTGATGTTGTTTCCCAATG 374
QY 503 CTTTAGAGGAGTTGTTTAAACCTTTTAAAGCCCACTTTTGTGCGGAGCGGTCAACC 562
Db 375 CTGTATAGAGAGTTATTGAAACCCCTTTAATGATGAAAAAGTAATGGCCACACAGTGTCCAG 434
QY 563 TTAATGTCAGAAATAGACAAACCAATCTCTTAACACGCTTGCAGATATATCCCTATGATA 622
Db 435 TGAACATTCGTAATAGAAATGATTAATTAACAAAACTAATTTGATATGCGTTATGACA 494
QY 623 ATGCTTTTGGCGTTGAACGAGCTGCCCAATTCGTTTACAGTAAATATATCTTTGTTGCTCAG 682
Db 495 ATGCGTTCCGTTGGAGCGTGCAGCAGTCCGTAACAGGAAATGTTCTTTGTTGATG 554
QY 683 GTCCGCTTACGTTTACAGACGCGAGGTGTTTCCCTAACATAGATAGATACATCAAC 742
Db 555 GCGCGTTAAGTTGTTATCGTAGAAGATAATACTGAAAAATTTAGAACATTAATGAAAGTC 614
QY 743 AGACCTTCTGCGGTATTCTCTAAGTATTGGTGAACAGGCTGTTGACCACTATGCAA 802
Db 615 AGATGTTCTTGGTGAAGAGGTGCGAGTTGGAGATGATAGATGCTTAATAATATGCTA 674
QY 803 CTGATTTAGAAAGACTGTTTATCAATCCACTGCTAAATGTTTATACAGATGTTCTGAC 862
Db 675 TTTTGAAGGAAACACAGTTTATCAATCCACTGCTCGATGTTTACTGATGCTCCAATA 734
QY 863 AGATGCTCTACTTACTTGAAGCAGCAAAACCGCTGGAACAGTCTCTTTTAGAGAGTCCA 922
Db 735 CATTAACCAATTTCTTAACAGCAACTAGTGTGAACAAAGTCAATTTTGTAGAGAAAGTT 794
QY 923 TTAATCTGTTAAGAAAAATCATGAACAAATCTTTTGTAGCCCTATGACCACTACTTGAGG 982
Db 795 TAATTTCACTTGGCATTTGTTATGAAAAACCAATGTTCTTTGTTGGACAATTTTCGAAA 854
QY 983 TGTCTATGTTTATGATGCTTTTATTTCTGTTGTTGTTGTTTGTAGGCAATGTCAGAG 1042
Db 855 TATCGTTATGGAATTTTATTTGGGCTTTCCCTACTTCTAAGTATTAATTTCTCAAGGCAAGTC 914
QY 1043 AATTTGATTCGCTAGGCTTTTACCTTTCTGTCGATTAATCTTCAATGTTGCGCTGTGTC 1102
Db 915 ATGATGTTAATTTTGGCTGTTTATTTATTTTGGGTTATATTTTCTAGCTGTATATGCTA 974
QY 1103 GGAACATTTCAATACATGCTTAAGCACCCGCTGCTCTTCTTCTTCTTCTCGTTTATGGGG 1162
Db 975 GAAATGATTTTATCTATTATAAACAATCCCTTACTTCTTACGCGCCATATATGAA 1034
QY 1163 TGCCTGATTTTGTCTCCTCAGCCCTTGAATTAATTTCTTTTACTATTAATAAGT 1222
Db 1035 TTCTCCATGTTAGCACTATTACCTATACGCTTTTATGCTTTTACTAACTAATAATCTA 1094
QY 1223 CTGACTGGGGAACACCACTAAAAAATTAATAT 1252
Db 1095 ATGCTTGGGGAACAGCTTAATTACGTAAT 1124

RESULT 14
US-09-902-939-2
; Sequence 2, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHayza
; APPLICANT: William Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/0H020-US0
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0

Query Match 4.9%; Score 61.6; DB 14; Length 4194;
Best Local Similarity 49.1%; Pred. No. 1.8e-06;
Matches 222; Conservative 0; Mismatches 224; Indels 6; Gaps 2;
474 TGATCGAGATATTATCTACCTGATGCGCTTTAGAGAGATGTTTAAACCTT---TAA 530

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2003, 09:01:59 ; Search time 2968 seconds
(without alignments)
10268.805 Million cell updates/sec

Title: US-09-469-200D-1
Perfect score: 1254
Sequence: 1 atgagaacattaaaaaacct.....cacgtataaaattattataa 1254

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```

ESP:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hct:*
9: gb_est1:*
10: gb_est2:*
11: gb_hct:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pbg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	61.6	4.9	2891	11	AK079729	AK079729 Mus muscu
c 2	50.6	4.0	1101	29	CNS017WW	AL108602 Drosophi
3	49.4	3.9	809	14	CA324064	CA324064 UI-M-FY0-
4	49.4	3.9	4151	11	AK028582	AK028582 Mus muscu

5	48.4	3.9	1180	13	BX436369	BX436369
6	48	3.8	712	13	BX416727	BX416727
7	46.8	3.7	712	28	BH980141	BH980141
8	46.8	3.7	919	29	CNS01MCA	BH980141 Odg45g12.
c 9	46.4	3.7	1702	10	BG165669	AL150747 Anopheles
10	46.2	3.7	1124	28	B12074	BG165669 602345171
c 11	45.8	3.7	976	29	CNS001Q2	B12074 R2549-r7-2
c 12	45	3.6	714	13	BX415371	AL075095 Drosophi
c 13	45	3.6	939	29	CNS00CNG	BX415371 BX415371
c 14	44.6	3.6	1176	13	BX333445	AL059400 Drosophi
c 15	44.4	3.5	1101	29	CNS00LOO	BX333445 BX333445
16	44.2	3.5	816	29	CNS02KCV	AL068607 Drosophi
17	44.2	3.5	1201	9	AL531744	AL201604 Tetraodon
18	44	3.5	485	13	BQ828135	AL531744 AL531744
c 19	44	3.5	1101	29	CNS00LFT	BQ828135 L16h1142
c 20	43.8	3.5	524	29	CNS007GT	AL078714 Drosophi
21	43.8	3.5	960	13	BX397886	AL067552 Drosophi
c 22	43.4	3.5	903	29	CNS0011C	BX397886 BX397886
c 23	43.4	3.5	1072	29	CNS007OB	AL074609 Drosophi
24	43.2	3.4	643	28	AZ282684	AL067418 Drosophi
c 25	43.2	3.4	654	10	BE889709	AZ282684 RRC1-23-1
c 26	43.2	3.4	945	29	CNS00IKW	BE889709 601512740
27	43.2	3.4	1100	13	BX381739	AL075212 Drosophi
28	43.2	3.4	1201	9	AL514129	BX381739 BX381739
29	43	3.4	976	13	BX347783	AL514129 AL514129
30	43	3.4	1000	29	CNS00COQ	BX347783 BX347783
c 31	42.8	3.4	864	29	CNS016M4	AL059446 Drosophi
32	42.6	3.4	399	29	CNS01609	AL106486 Drosophi
33	42.4	3.4	473	29	BX238763	AL106131 Drosophi
34	42.4	3.4	907	29	CNS021J4	BX238763 Danio rer
35	42.4	3.4	1249	29	BZ566650	AL176953 Tetraodon
c 36	42.2	3.4	498	9	AL574975	BZ566650 pacs2-164
c 37	42.2	3.4	530	13	BX425377	AL574975 AL574975
c 38	42.2	3.4	883	29	CC261953	BX425377 BX425377
c 39	42.2	3.4	913	28	BH155504	CC261953 CH261-184
c 40	42.2	3.4	915	28	BH139902	BH155504 ENTRS33TR
c 41	42.2	3.4	940	9	AL581799	BH139902 ENTS49TR
c 42	42.2	3.4	957	28	AZ668597	AL581799 AL581799
c 43	42.2	3.4	1472	10	BG421340	AZ668597 ENTG03RTF
c 44	42	3.3	864	29	CNS0605G	BG421340 602451464
c 45	42	3.3	926	28	BH943155	AL076642 T7 end of
						BH943155 odd64a12.

ALIGNMENTS

RESULT 1
AK079729
LOCUS
DEFINITION
AK079729 2891 bp mRNA linear HTC 05-DEC-2002
Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched library, clone:A43004J01 product:hyaluronan synthase 2, full insert sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

1
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED
11042159


```

CNS017WW/c
LOCUS      CNS017WW      1101 bp      DNA      linear      GSS      26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN37D07 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION  AL108602
VERSION    AL108602.1  GI:5628906
KEYWORDS   GSS.
SOURCE     Drosophila melanogaster (fruit fly)
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE  1 (bases 1 to 1101)
AUTHORS    Genoscope.
TITLE      Direct Submission
JOURNAL    Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
            - web : www.genoscope.cns.fr)
COMMENT    Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Payan. It has been constructed in the vector
            pBelobAC11.

FEATURES             Location/Qualifiers
     source           1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN37D07"
                     /clone_lib="DrosBAC"
                     /plasmid="pBelobAC11"
                     /note="end : SP6"

BASE COUNT  533 a 179 c 108 g 125 t 156 others
ORIGIN
Query Match      4.04; Score 50.6; DB 29; Length 1101;
Best Local Similarity 42.4%; Pred. No. 1.1;
Matches 131; Conservative 22; Mismatches 156; Indels 0; Gaps 0;

QY  907  TTCCTTAGAGTCCATATTCTCTTAGAAATCATGACAAATCCCTTTAGCCCTA 966
      || || || || || || || || || || || || || || || || || ||
Db  487  TTCTTCTGTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 428
      || || || || || || || || || || || || || || || || || ||

QY  967  TGGACCATCTGAGGTGCTATGTTATGATGCTTTTATCTGCTGGATTTCTTT 1026
      || || || || || || || || || || || || || || || || || ||
Db  427  TGTAGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 368
      || || || || || || || || || || || || || || || || || ||

QY  1027 GTAGGCAATGTCAGAGAAATTTGATTTGGCTCAGGTTTATAGCCCTTCTGGTATTCTTC 1086
      || || || || || || || || || || || || || || || || || ||
Db  367  TTWTCTTATCTCATATTTTGTGTYTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 308
      || || || || || || || || || || || || || || || || || ||

QY  1087 ATGTGTCCCTGTCGGAACATTCATTACATGCTTAACACCCCGCTGCTCTCTGTTA 1146
      || || || || || || || || || || || || || || || || || ||
Db  307  TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 248
      || || || || || || || || || || || || || || || || || ||

QY  1147 TCTCCGTTTATGGGTGCTGATTTGCTTCTCTACACCCCTTCAATATATATCTCTT 1206
      || || || || || || || || || || || || || || || || || ||
Db  247  TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 188
      || || || || || || || || || || || || || || || || || ||

QY  1207 TTACTATT 1215
      || || || || || || || || || || || || || || || || || ||
Db  187  TTGTGTTATT 179
      || || || || || || || || || || || || || || || || || ||

RESULT 3
CA324064
LOCUS
DEFINITION UI-M-FY0-cso-g-16-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
            IMAGE: 6822281.5', mRNA sequence.

```

```

CA324064
CA324064.1  GI:24542162
EST.
KEYWORDS   Mus musculus (house mouse)
SOURCE     Mus musculus
ORGANISM   Mus musculus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 809)
REFERENCE  NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished
            Contact: Robert Strausberg, Ph.D.
            Email: cgraphs@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
Seq primer: pyX-5.
            Location/Qualifiers
             1..809
             /organism="Mus musculus"
             /mol_type="mRNA"
             /strain="C57BL/6"
             /db_xref="taxon:10090"
             /clone="IMAGE: 6822281"
             /tissue_type="whole brain"
             /dev_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"
             /lab_host="DHL0B (11 phage resistant)"
             /clone_lib="NIH_BMAP_FY0"
             /note="Organ: Brain; Vector: pyX- Asc; Site: 1: EcoR I;
            Site: 2: Not I; The library was constructed according
            Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
            1996. Denatured RNA was size fractionated on a 1% agarose
            gel. First strand cDNA synthesis was primed with oligo-dT
            primer containing a Not I site. Double strand cDNA was
            size selected according to mRNA size fraction, ligated
            with EcoR I adaptor, digested with NotI and then cloned
            directionally into pyX-Asc vector. The library tag
            sequence located between the Not I site and the polyA tail
            is ACGGAGACAG. This library was created for the University
            Iowa Brain Anatomy Project (BMAP). Gene Discovery in the
            Developing Mouse Nervous System", supported by National
            Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
            program coordinator."
BASE COUNT  175 a 218 c 193 g 221 t 2 others
ORIGIN
Query Match      3.9%; Score 49.4; DB 14; Length 809;
Best Local Similarity 49.2%; Pred. No. 1.9;
Matches 159; Conservative 0; Mismatches 161; Indels 3; Gaps 1;

QY  600  CTTGACAGATATTCCTATGATATGCTTTGGCGTTGACAGAGCTGCCCATCCGTTAC 659
      || || || || || || || || || || || || || || || || || ||
Db  119  CCTGCGAGTGTGAGTAGTATGATGCTTTTCAAGTGGAGCGGGCTGCCACTCTACTT 178
      || || || || || || || || || || || || || || || || || ||

QY  660  AGGTAATATCCTTGTGTTCTCAGGTCGCGTTAGCGTTTACACGCGAGGTGGTTGTTC 719
      || || || || || || || || || || || || || || || || || ||
Db  179  TGGCTGTGCAATCTATAGTGGCGCTTTGGGCATGATCCGACACCCCTCTCAGCA 238
      || || || || || || || || || || || || || || || || || ||

QY  720  TAACATAGATAGATACATCAACGACCTCTCTGGGTATTCCTGTGAAGTATTTGGTGATGA 779
      || || || || || || || || || || || || || || || || || ||
Db  239  GTTCTCGAGGATTTGGTACCATCAGAAGTCTCTAGGCAGCAAGTCAGCTTTGGGATCA 298
      || || || || || || || || || || || || || || || || || ||

QY  780  CAGGTCTTGACCAACTATGCACTGATTAGGAAA---GACTGTTTATCAATCACTCTGC 836
      || || || || || || || || || || || || || || || || || ||
Db  299  TCGGCACTTACCAACCGAGTCTCTGAGTCTTGGCTACCGGACTTAAGTATACAGACGCTC 358
      || || || || || || || || || || || || || || || || || ||

QY  837  TAAATGTTATACAGATGTTCTCTGACAAAGATGTTACTTACTTGAAGCAGCAAAACCGCTG 896
      || || || || || || || || || || || || || || || || || ||

```



```

QY 1037 TCAGAGATTGATGGCTCAGGGTTTATGACCTTCTGTGATTATCTTCATTGTTGCC 1096
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 497 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 556
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1097 TGTGTCGACATTCATACATGCTTAAGCACCCTGCTCTCTGTTATCTCGCTTT 1156
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 557 KKKAKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 616
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1157 ATGGGTCGTGCTATTTGTTGCTTACAGCCCTGAAATATATATCTCTTTTACIATA 1216
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 617 KKKAKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 676
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1217 GAAATGCTGACTGGG 1232
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 677 KKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 692
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
BH980141
LOCUS
DEFINITION
odg45g12.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
sequence.
ACCESSION
BH980141
VERSION
BH980141.1 GI:23510861
KEYWORDS
GSS.
SOURCE
Brassica oleracea
ORGANISM
Brassica oleracea
REFERENCE
1 (bases 1 to 712)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Brassica.
AUTHORS
Delehaunty,K., Fellw.,G., Fulton,L., McCombie,W.R., Miner,T., Nash
,W., Rabinowicz,P.D. and Wilson,R.K.
TITLE
Whole genome shotgun reads from Brassica oleracea
JOURNAL
Unpublished
COMMENT
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: odg45 row: g column: 12
Seq primer: -21UpPOT forward
Class: shotgun
High quality sequence start: 17
High quality sequence stop: 551.
FEATURES
source
1..712
Location/Qualifiers
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea002"
/note="Vector: pOTw13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea T01000DH3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."
BASE COUNT 232 a 66 c 75 g 338 t 1 others
ORIGIN
Query Match 3.7%; Score 46.8; DB 28; Length 712;
Best Local Similarity 52.5%; Pred. No. 6.9;
Matches 125; Conservative 0; Mismatches 112; Indels 1; Gaps 1;

QY 922 ATATTCTGTTAAGAAATCATGACAACTCTTTGTAGCCCTATGACCACTGAG 981
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 147 ATTATTTTATTAATAATATTTTCATATTTTGTGAATTTTATTTTGTGATTT 206
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 982 GT-GCTATGTTAAGATGCTGTTTATTCGTGTTGGAATTCCTTGTAGCAATGTCAG 1040
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 207 TTAGTGACTTTTATGATTTTTCATTTTATGATTTTATGATTTTATTTCTTATTTATA 266
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1041 AGAATTTGATGGCTCAGGGTTTACGCTTCTGTTGATATATCTCAATGTCGCCCTGTG 1100
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 267 TTATTTTATTAATAATAATATTTTCTGCGAATATTTGATTTTGTATATT 326
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1101 TCGGAACATTCATACATGCTTAAGCACCCTGCTCTCTGTTATCTCGCTTTAT 1158
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 327 TAGTGACTTTATGATTTGTTTGAATTTTATGAATTTTATATATTTTCTTTT 384
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
CNS01MCA
LOCUS
DEFINITION
Anopheles gambiae GSS T7 end of clone 21G21 of Notre Dame library
from strain PEST of Anopheles gambiae (African malaria mosquito),
genomic survey sequence.
ACCESSION
AL150747
VERSION
AL150747.1 GI:7011226
KEYWORDS
GSS.
SOURCE
Anopheles gambiae (African malaria mosquito)
ORGANISM
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
REFERENCE
1 (bases 1 to 919)
Genoscope.
AUTHORS
Direct Submission
JOURNAL
Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
REFERENCE
2 (bases 1 to 919)
Roth,C.W., Brey,P.T., Ke,Z., Collins,F.H. and Welissenbach,J.
Submitted (16-FEB-2000) BML, Institut Pasteur, 25, rue du Dr.
Roux, Paris 75015, France
This clone is from an A. gambiae BAC library provided by F.H.
Collins and sequenced by Genoscope in collaboration with the
Laboratory of Biochem. and Biol. Molec. of Insects, Institut
Pasteur.
FEATURES
source
1..919
Location/Qualifiers
/organism="Anopheles gambiae"
/mol_type="genomic DNA"
/strain="PEST"
/db_xref="taxon:7165"
/clone="21G21"
/clone_lib="Notredame1"
/note="end : T7"
BASE COUNT 192 a 165 c 153 g 359 t 50 others
ORIGIN
Query Match 3.7%; Score 46.8; DB 29; Length 919;
Best Local Similarity 42.9%; Pred. No. 7.2;
Matches 133; Conservative 26; Mismatches 148; Indels 3; Gaps 1;

QY 907 TCTCTAGAGCTCCATATTTCTGTAGAAATCATGACAACTCTTTGTAGCCCTA 966
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 TTTTCTTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 174
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 967 TGGACCATATCTGAGGTGCTATGTTTATGATGCTGTTTATCTGTGGGATTTCTTT 1026
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 WTATAACACTCWTCAATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 234
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1027 GTAGGCAATGTCAGAAATTTGATGGCTCAGGCTTTTAGCCCTTCTCGGTGATATCTTC 1086
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 TTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 291
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1087 ATGTTTGCCCTGCTCGGACATTCATACATGCTTAAGACCCCTGCTCTCTGTGTA 1146
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 292 ATTTTATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 351
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1147 TCTCCGTTTATGAGGGGCTGTCATTTGTTGCTCAGACCCCTTGAATATATATCTCTT 1206
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 352 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 411
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

QY	1207 TTACTATTA 1216
Db	412 TTTTAACTA 421
RESULT 9	
LOCUS	BG165669/c
DEFINITION	60234517F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:445353 5', mRNA sequence.
ACCESSION	BG165669
VERSION	BG165669.1 GI:12672372
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1. (bases 1 to 1702) NIH-MGC http://mgc.nci.nih.gov/.
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnlnl.gov plate: LLAM10248 row: m column: 18 High quality sequence stop: 542.
FEATURES	location/Qualifiers 1..1702 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:445353" /tissue_type="hypernephroma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_89" /note="Organ: kidney; Vector: pCMV-SPORT6; Site.1: NotI; Site.2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library." BASE COUNT 802 a 286 c 338 g 276 t ORIGIN
Query Match	3.7%; Score 46.4; DB 10; Length 1702;
Best Local Similarity	48.5%; Pred. No. 9.8;
Matches 128; Conservative	0; Mismatches 136; Indels 0; Gaps 0;
QY	951 TCCTTTGTAGCCCTATGGACATACTTCAGGTGCTCATGTATTATGATCGTTTATTC 1010
Db	1585 TTCCTTTTGCTCTTTGCTCCTTCTCTGTTTTGTTTGTGCGCTCTCTCTTTTTCCTTTC 1526
QY	1011 TGTGTTGATTCCTTTTAGCGAATGTCAGAGAATTGATGCTCAGGGTTTASCCTT 1070
Db	1525 TCTGTTCTTCTTTTGTGTTTCTTCTCTGTAATTCCTCTCTTCCTTCTGTTTCTTCT 1466
QY	1071 TCTGTTGAATTCCTCAATGTTGCCCTGTGCGGAACATCATACATGCTTACGACC 1130
Db	1465 CTGCTTCTTCTCTTCCTGTTTTCCTCTTCTCCGTTGTTGCTTTTGTCTTC 1406
QY	1131 GCTGCTCTTCTTGTATCTCGGTTTATGGGTCGTCGAATTTGTTGCTACAGCCCT 1190
Db	1405 TTCCTCCCTCTCTTCTTCTTCCCTCTCTCTTCTTCTTCTCTCATCTTCTTCTTGCCTTC 1346
QY	1191 GAATATATATTCCTTTTTACTAT 1214
Db	1345 TCTGCTCTTCTTCTTCTTCTTCTGTT 1322

```

RESULT 12
BX415371/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1. .714
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOCAP005YL10"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 303 a 139 c 131 g 97 t 44 others
ORIGIN

Query Match 3.6%; Score 45; DB 13; Length 714;
Best Local Similarity 40.9%; Pred. No. 17;
Matches 99; Conservative 27; Mismatches 116; Indels 0; Gaps 0;

QY 974 TACTTGAGGTGCTATGATTATGATGACTTGTTATCTGTGGTGGAATCTTTGTAGGCA 1033
||| : |||| :||: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
709 TACTAKGGTGGTGTATKYGKGTGTGGKGGKTTTGTGAGTTTGTGATTTTGTGTTGTTT 650
: : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 1034 ATGTACAGAGAAATTCATTGGCTCAGGTTTGTAGCCCTTCTCGTGATATCATCTGTTG 1093
: : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 649 TGGGGGKGYYTATTKTGKTGATNATNTATTTWKWBCBTGTTGTTGTGTTGTTT 590
: : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 1094 CCCTGTGTGGAGCAATTCATTACATTCATTAAGCACCGCGTGTCTCTGTTATCTCCGT 1153
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 589 KGGTYYTATTKGTDWTTTYYHKKWTTTGTGTTTGTGTTTGTGTTTGTGTTT 530
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 1154 TTATGGGGTGCTGCATTTGTTTGCCTACAGCCCTGAAATATATCTCTTTTACTA 1213
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 529 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 470
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
QY 1214 TT 1215
||
Db 469 TT 468

RESULT 13
CNS000CNG/c
LOCUS
DEFINITION
ACCESSION

```

VERSION	AL059400.1	GI:4946964
KEYWORDS	GSS.	
SOURCE	Drosophila melanogaster (fruit fly)	
ORGANISM	Drosophila melanogaster	
	Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Pflydoidea; Drosophilidae; Drosophila.	
REFERENCE	1 (bases 1 to 939)	
AUTHORS	Genoscope.	
TITLE	Direct Submission	
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)	
COMMENT	Web 1 : www.genoscope.cns.fr Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-96 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's Pi and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .	

[illegible]

RESULT 14	
BX333445/c	
LOCUS	BX333445 mRNA 1176 bp EST 01-MAY-2003
DEFINITION	Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DC025YN02 5'-PRIME, mRNA sequence.
ACCESSION	BX333445
VERSION	BX333445.1 GI:30308208
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 1176)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9051.r For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DC025DG01QP1
cluster=9051.r. Contact : Feng Liang Email : fliang@lifetech.com
URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DC025DG01QP1.
Location/Qualifiers
1..1176
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC025N02"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/culture_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 390 a 176 c 239 g 208 t 163 others
ORIGIN

```

Query Match      3.6%; Score 44.6; DB 13; Length 1176;
Best Local Similarity 34.3%; Pred. No.23;
Matches 113; Conservative 51; Mismatches 165; Indels 0; Gaps 0;
```

QY	900 CAAGTCCTCTTAGAGAGTCATATTCTGTGAAGAATAATCATGACAAATCCTTTTGGT 959
Db	: ::: :: :
QY	476 C AAGGTTATTTTKWKCOWTTTTTYIKWTWTCACAAKTATACATAYAACCCCTTT 417
Db	: ::: :: :
QY	960 AGCCCTATGGACCATACTTGAGGTCGTCAATGTTTAATGATCGTTGTTATTCGTGGTGA 1019
Db	416 TGCSTTTTTKGRACCCTYGATGCTKTYTCATWYTGGTGCATGCTGCTWTTTTCGCG 357
QY	1020 TTCTTTTGTAGCAATGTCAGAGAATTTGATTTGGCTCAGGGTTTATAGCCTTTCGGTGAT 1079
Db	: : ::: ::: ::: ::: ::: ::: ::: ::: :::
QY	356 ITWCTTTGTAAITTTTAAIKYIVCYKTRAKTTTTTYWKNTTTTTTTTKWATCCVKTT 297
Db	: ::: :: :
QY	1080 TATCTTCAATGTGCCCTGTGTCGGAACATTCATTACATGCTTAAGCACCCGCTGTCCTT 1139
Db	: : : : : : : : : : : : : : : :
QY	296 ITTYFTKCTWTTTTTTTANKCKCKTKTKCKKKITTYCKTKTKYCWKCKGCKVKTT 237
Db	: ::: :: :
QY	1140 CTGTGTTATCTCCGTTTATGGGGTGCTGCAATTTGTTTGTOCTACAGCCCTTGAATATA 1199
Db	: : ::: ::: ::: ::: ::: ::: ::: ::: :::
QY	236 TTCTTTTYYKIIITTTTTTTTTTTTTTVCATATKTTTTCWKKWTTTTTYKYKTTTTTTTT 177
Db	: ::: :: :
QY	1200 TTCCTCTTTTACTATAGAAATGCTGACT 1228
Db	: ::: :: :
QY	176 CCTTTTITTYTTCOCCTTGKTTACT 148
Db	: ::: :: :

RESULT 15
CNS00L00/c
LOCUS CNS00L00 1101 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TEf3 end of BAC:
BACR32023 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL068607
VERSION AL068607.1 GI:4958689
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2003, 10:30:08 ; Search time 361 Seconds
(without alignments)
3118.187 Million cell updates/sec

Title: US-09-469-200d-2
Perfect score: 2145
Sequence: 1 MRTLKILITVAFSIFWLL.....KLYSLFIRNADWGRKKLL 417

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2.1/USPO1/US09469200/runat_02102003_090347_22770/app-query.fasta.1.583
-DB=N_Geneseq_19Jun03 -QPM=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09469200 @CGN.1.1.312 @runat_02102003_090347_22770 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2145	100.0	1254	20	AA558841	Streptococcus equi
2	1586.5	74.0	1335	24	ABN67167	Streptococcus poly
3	1586.5	74.0	1512	25	AB857951	S. pyogenes hyalur
4	481.5	22.4	7193	21	AAA14914	DNA encoding nodul
c 5	463	21.6	534720	19	AAV30458	Rhizobium species
c 6	463	21.6	536165	19	AAV30459	Rhizobium species
7	447.5	20.9	1662	25	ACC51023	Human bladder canc
8	447.5	20.9	1662	25	ABX76415	Lung cancer-associ
9	447.5	20.9	1662	25	ABX76414	Lung cancer-associ
10	440	20.5	1665	19	AAV18822	Murine hyaluronan
11	440	20.5	1665	21	AA288201	Mouse hyaluronate
12	440	20.5	1665	25	AB276736	Mouse hyaluronan s
c 13	439	20.5	1653	24	ABL41013	Murine hyaluronan
14	439	20.5	1659	25	ABZ76735	Mouse hyaluronan s
15	439	20.5	2947	19	AAV18821	Murine hyaluronan
16	439	20.5	2948	24	AB199697	Mouse ischaemic co
17	433.5	20.2	3003	24	ABK84373	Human cDNA differe
18	433.5	20.2	4018	22	AAU26639	Human breast cance
19	428.5	20.0	2890	21	AA288200	Mouse hyaluronate
20	370.5	17.3	2117	18	AA296713	Human hyaluronate
21	367	17.1	1752	25	AB276734	Mouse hyaluronan s
22	367	17.1	2102	18	AA291655	Mouse hyaluronate
23	367	17.1	2102	20	AA210862	Hyaluronate synth
24	367	17.1	2102	21	AAA39987	Murine HAS1 cDNA
25	367	17.1	2102	21	AA288199	Mouse hyaluronate s
26	366.5	17.1	2116	18	AA299541	Human hyaluronan s
27	336.5	15.7	1740	20	AA558846	Chlorella virus PB
c 28	265.5	12.4	265118	22	AAH41227	Pyrococcus abyssi
c 29	265.5	12.4	349980	22	AAH41226	Pyrococcus abyssi
30	262.5	12.2	2365589	24	ABA90521	Genomic sequence o
31	245.5	11.4	1239	22	AA289348	Staphylococcus aur
32	245.5	11.4	1239	22	AA289349	Staphylococcus aur
33	245.5	11.4	5648	22	AA289353	Staphylococcus aur
34	242.5	11.3	4951	17	AA205848	Polycistronic cps
35	242.5	11.3	4951	20	AA211802	Nucleotide sequenc
c 36	242.5	11.3	7430	18	AAV74571	Staphylococcus aur
37	234.5	10.9	1257	24	ABN91751	Staphylococcus epi
38	234.5	10.9	4500	21	AA287998	Nucleotide sequenc
39	233.5	10.9	570	24	ABN96952	Gene #3450 used to
40	226	10.5	1266	24	ABN67166	Streptococcus poly
41	226	10.5	1266	24	ABN70371	Streptococcus poly
c 42	226	10.5	215561	24	ABN71527	Streptococcus poly
43	198.5	9.3	1203	24	ABN71063	Streptococcus poly
c 44	195.5	9.1	13884	20	AA213278	Enterococcus faeca
c 45	195.5	9.1	13884	24	AB599073	Enterococcus faeca

ALIGNMENTS

RESULT 1
AA558841
ID AA558841 standard; DNA; 1254 BP.
AC AA558841;
XX
XX 16-AUG-1999 (first entry)
DT
XX Streptococcus equisimilis hyaluronate synthase seHAS DNA.
DE Hyaluronate synthase; seHAS; hyaluronic acid; hyaluronan; ss.
XX Streptococcus equisimilis.
OS Streptococcus equisimilis.
XX
XX Key Location/Qualifiers
FT primer_bind complement (316..337)
FT /*tag= a

FT primer_bind /note= "primer sel, given in AAX58842"
 475...494
 FT /*tag= b
 FT /note= "primer seep1, given in AAX58844"
 1031...1050
 FT /*tag= c
 FT /note= "primer se2, given in AAX58843"
 1228...1244
 FT /*tag= d
 FT /note= "primer seep2, given in AAX58845"
 XX
 PN W09923227-A2.
 XX
 PD 14-MAY-1999.
 XX
 XF 30-OCT-1998; 98WO-US23153.
 XX
 PR 26-OCT-1998; 98US-0178851.
 PR 31-OCT-1997; 97US-0064435.
 XX
 PA (OKLA) UNIV OKLAHOMA STATE.
 XX
 PI DeAngelis P, Kumari K, Weigel PH;
 XX
 DR WPI; 1999-337486/28.
 DR P-PSDB; AAY06206.
 XX

XX Nucleic acid encoding hyaluronate synthase for production of
 PT hyaluronic acid with controlled molecular weight and targeting
 PT specificity
 XX
 PS Claim 3; Page 118; 125pp; English.

XX This DNA sequence codes for a hyaluronate synthase, termed seHAS
 CC (see AAY06206), of group C Streptococcus equisimilis. To isolate
 CC the DNA, genomic DNA was subjected to PCR amplification using
 CC degenerate primers (see AAX58847-52) based on known related
 CC sequences, and PCR products were used as probes to screen a DNA
 CC library. Expression of the seHAS gene correlates with virulence
 CC of streptococcal Group A and Group C strains, by providing a means
 CC of escaping phagocytosis and immune surveillance. The invention
 CC also provides recombinant vectors containing the isolated DNA, and
 CC prokaryotic or eukaryotic host cells which produce seHAS and its
 CC hyaluronic acid product, particularly a product with modified
 CC structure or molecular size. The hyaluronic acid produced this way
 CC is purer than that produced by conventional methods. The seHAS
 CC gene also provides a new probe to assess the potential of bacterial
 CC specimens to produce hyaluronic acid.

XX SQ Sequence 1254 BP; 335 A; 240 C; 261 G; 418 T; 0 other;

Alignment Scores:
 Pred. No.: 3-24e-214 Length: 1254
 Score: 2145.00 Matches: 417
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 20 Gaps: 0

US-09-469-200D-2 (1-417) x AAX58841 (1-1254)

OY 1 MetArgThrLeuLysAsnLeuIleThrValValAlaPheSerIlePheTrpValLeu 20
 Db 1 ATGAGACATTAAAAACCTCATAACTGTGTGGCCITTAGTATTTTGGTACTGTG 60
 OY 21 IleTyrValAsnValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeu 40
 Db 61 ATTACCTCAAGTTTATCTCTTGTGTGCTAAAGAGCTGTCAATTTATGGCTTTTG 120
 OY 41 LeuIleAlaTyrLeuLeuValLysMetSerLeuSerPhePheTyrIlePheLysGly 60
 Db 121 CTGATAGCTACCTATTAGTCAAAATGCTTCCTTTCCTTTTACAAAGCCATTAAAGGA 180

OY 61 ArgAlaGlyGlnTyrIleValAlaAlaIleLeuProSerTyrAsnGluAspAlaGluSer 80
 Db 181 AGGCGTGGCAATATAGGTTGCAGCCATTATTCCTCTTATAACGAAGATCGTAGTCA 240
 OY 81 LeuLeuGluThrLeuLysSerValGlnGlnThrTyrProLeuAlaGluIleTyrVal 100
 Db 241 TTGCTAGAGACCTTAAAGATGTTTCAGCAGCAAACTATCCCTAGCAAAATTTATGTT 300
 OY 101 ValAspAspGlySerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgAsp 120
 Db 301 GTTGACGATGAAGTCTGATGAGACAGGTATTAAAGCGCATTTGAAGCATTTGCGTGAC 360
 OY 121 ThrGlyAspLeuSerSerAsnValIleValHisArgSerGluLysAsnGlnGlyArg 140
 Db 361 ACTGCTGACCTATCAAGCAATGCTATGTTTCATCGGTACAGAAAAATCAAGGAAGCGT 420
 OY 141 HisAlaGlnAlaTrpAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSer 160
 Db 421 CATGCACAGGCTGGGCGCTTTGAAGCATCAGACGCTGATGCTTTTACCGTTGACTCA 480
 OY 161 AspThrTyrIleTyrProAspAlaLeuGluGluLeuLysThrPheAsnAspProThr 180
 Db 481 GATACTATATCTACCTGATGCTTTAGAGAGATTGTTAAAAACCTTTAATGACCAACT 540
 OY 181 ValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuThrArg 200
 Db 541 GTTTTGTCTGGCAGGCTCCTTAAATGTCAGAAATAGACAAACCAATCTTTAACACGC 600
 OY 201 LeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThr 220
 Db 601 TTGACAGATATTGCGCTATGATGCTTTTGGCGTTGAAGAGCTGCCCAATCCGTTACA 660
 OY 221 GlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgGluValValPro 240
 Db 661 GGTAATATCTTTGCTCGCTGCTTAGCGTTTACAGACGGAGTGGTTGCTCT 720
 OY 241 AsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAsp 260
 Db 721 AACATAGATAGATACATCAACAGACCTTCCTGGGTATTCCTGTAGTATTTGGTGAAC 780
 OY 261 ArgCysLeuThrAsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLys 280
 Db 781 AGTGTCTTGACCAACTATGCACTGATTTAGAAAGACTGTTTATCAATCCACTGCTAAA 840
 OY 281 CysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrpAsn 300
 Db 841 TGTATTACAGATGTTCTGACAGATGCTACTTACTTGAAGCAGAAAACCGCTGGAAC 900
 OY 301 LysSerPhePheArgGluSerIleLeuSerValLysLysIleMetAsnAsnPropheVal 320
 Db 901 AAGTCTCTTTTAGAGAGTCCATTATTTCTGTTAAGAAAATCATGNACATCTTTTGTGA 960
 OY 321 AlaLeuTrpThrIleLeuGluValSerMetPheMetMetLeuValTyrSerValValAsp 340
 Db 961 GCCTATGGACCACTTACGTGTCTATGTTTATGCTTGTATTCTGTTGCTGGAT 1020
 OY 341 PhePheValGlyAsnValArgGluPheAspTrpLeuArgValLeuAlaPheLeuValIle 360
 Db 1021 TTTCTTTGAGCAATGTCAGAAATTTGATTTGGCTCAGGGTTTTAGCCTTTCTGTGTGAT 1080
 OY 361 IlePheIleValAlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPhe 380
 Db 1081 ATCTTCATTTGTGCTGCTGTGCGGAACATTATCATGCTTAAGCAGCCGCTGCTTTC 1140
 OY 381 LeuLeuSerProPheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyr 400
 Db 1141 TTGTTATCTCGGTTTATGGGGTGTCTCATTTGTTTGTCTTACAGCCCTTGAATATAT 1200
 OY 401 SerLeuPheThrIleArgAsnAlaAspTrpGlyThrArgLysLysLeuLeu 417
 Db 1201 TCTCTTTTACTATTAGAAATGCTGACTGGGACACACGTAATAAATTTATTA 1251

RESULT 2

Claim 1; Fig 1; 47pp; English.

CC The present sequence encodes nodulation efficiency factors. The present
 CC sequence is a 7.2 kb EcoRI/BamHI fragment from *Sinorhizobium meliloti*
 CC USDA 1170, that gives S. meliloti strain N8G 185 the ability to nodulate
 CC at least 50% of inoculated *Medicago lacinata* (cut-leaf medic) plants
 CC within 10 days of inoculation. The fragment contains nod genes nodA,
 CC nodB, nodC, nodD, nodJ. The nodulation factors are used to
 CC increase the nodulation efficiency of *Sinorhizobium* for *Medicago*
 CC *lacinata*. The nodulation factors are also used to improve nitrogen
 CC fixation in legumes.
 XX

SQ Sequence 7193 BP; 1434 A; 2094 C; 2132 G; 1532 T; 1 other;

Alignment Scores:

Pred. No.: 2,47e-39 Length: 7193
 Score: 481.50 Matches: 133
 Percent Similarity: 49.50% Conservative: 65
 Best Local Similarity: 33.25% Mismatches: 169
 Query Match: 22.45% Indels: 33
 DB: 21 Gaps: 10

US-09-469-200D-2 (1-417) x AAAL4914 (1-7193)

QY 33 SerLeuSerIleTyrGlyPheLeuLeuIleAlaTyrLeuLeuValLysMetSerLeuSer 52
 DB 1811 GCATCTGATTTACGGCGTCTCTTGCGCGCTAC-----AGGACATGCAA 1858
 QY 53 PhePheTyr-----LysProPheLysGlyArg 61
 DB 1859 GTCTTATATACTCGCGCGATAAACGGTCCACCGCTGTCGGCAGAACCGATCGAGACCGC 1918
 QY 62 AlaGlyGlnTyrLysValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSerLeu 81
 DB 1919 GCGCTCTGCGCCGCGTGGACGTTATGTCGCCAGTTTCATCCAGGATTCATCAGGACCGCATCTC 1978
 QY 82 LeuGluThrLeuLysSerValGlnGlnThrTyrPro---LeuAlaGluIleTyrVal 100
 DB 1979 TCGCGCTCGCTCGGTCATTCACACACGAGATTCCTGGAGACTGCGAGTCTATGTC 2038
 QY 101 ValAspAspGlySerAlaAspGluThrGlyIleLysArgIleGluAsp---TyrValArg 119
 DB 2039 GTTGATGATGGCTCTCGGAACCGCGAGCAATCGTCGTTACAGGATTTCTTATTCGCGC 2098
 QY 120 AspThrGlyAspLeuSerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLys 139
 DB 2099 GAT-----CCGAGGTTACCTTCATCTCTCCACAGAGACGTCGGAAG 2143
 QY 140 ArgHisAlaGlnAlaTTPalaPheGluArgSerAspAlaAspValPheLeuThrValAsp 159
 DB 2144 CGAAGCGCGAGATTCGCGGATAGGTCAATCTCTGGGAACTGGTGTGTAATGTGAC 2203
 QY 160 SerAspThrTyrIleTyrProAspAlaLeuGluLeuLysThrPheAsnAspPro 179
 DB 2204 TCAGACAGCAGANTGCTTTCGATGTCGTCCTCCAAAGCTTGCTCGAAGATGCGAAATCCA 2263
 QY 180 ThrValPheAlaIleThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThr 199
 DB 2264 GAGGTGTGGCGGCATGGTCACTACAGCTAGCAATCGGAGTGCACCTCGCTGAGC 2323
 QY 200 ArgLeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerVal 219
 DB 2324 AGATTGATCGACATGGATTTGGCTTGCTGCAAGAGAGCGCGCGCACAGGCTCGC 2383
 QY 220 ThrGlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgGluValValVal 239
 DB 2384 TTCGGTGGCGTTATGTGTGTCGCGGCCATGTGCTATGCTACCGTCTCGCGCTCGCT 2443
 QY 240 ProAsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAsp 259
 DB 2444 GCGCTGCTGACAGTACGAACGCGCGTGTTCGCGGTAAACACGCACTTCGCGTGAG 2503
 QY 260 AspArgCysLeuThrAsnTyrAlaThrAspLeuGly---LysThrValTyrGlnSerThr 278
 DB 2504 GATCGTCTCTGAGGATCTCATGTTGAGGCGAGGCTTAGACCGAGATGTGTCAGAC 2563

QY 279 AlaLysCysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArg 298
 DB 2564 GCCATAGTGGCAGCGCTGTTCCGGATAGGCTGAAGCGGTATCTACGCCAACAACTCGCT 2623
 QY 299 TrpAsnLysSerPheArgGluSerIleIleSerValLysLysIleMetAsn---Asn 317
 DB 2624 TGGCAGCAGACGCTTCGGGACAGCTTTTGGCGCTCCCTCTGTGGGCGGCGCTCAAC 2683
 QY 318 ProPheValAlaLeuTyrTrpIleLeuGluValSerMetPheMetMetLeuValTyrSer 337
 DB 2684 CGTATCTCATCTCATGATCGGTCGGCAGACTATCGGCCCATTTGTCTGCCCTCTCG 2743
 QY 338 ValValAspPhePheValGlyAsnValArgGluPheAspTrpLeuArgValLeuAlaPhe 357
 DB 2744 ATAGTGAGC-----GGACTTGGCGATTTTCATCATGACGCCACACACTCGCTGG 2791
 QY 358 LeuValIleIlePheIleValAlaLeu-----CysArgAsnIleHisTyrMet 373
 DB 2792 TGGACATTTGATGATTCATCCATCGCGCCATGATACGTCGACCGCTGGCATGGAAT 2851
 QY 374 LeuLysHisProLeuSerPheLeuLeuSerProPheTyrGlyValLeuHisLeuPheVal 393
 DB 2852 GCTCGCCAA---CTTAGATTCTTGGTACGTTCTACACACACCCCATCAACCTTTCTC 2908
 QY 394 LeuGlnProLeuLysLeuTyrSerLeuPheThrIleArgAsnAlaAspTrpGlyThrArg 413
 DB 2909 CTAATTCGTTGAAAGCCTATGCGTTGTATCATTTGCCAACAGGAGCTGTTGTCACGC 2968
 RESULT 5
 AAV30458/C
 ID AAV30458 standard; DNA; 534720 BP.
 AC AAV30458;
 XX
 DT 14-OCT-1998 (first entry)
 XX Rhizobium species plasmid pNGR234a.
 DE
 XX
 KW Symbolism; open reading frame; ORF; plasmid; vector; transportation;
 KW degradation; metabolism; host range; nitrogen fixation; nodulation;
 KW legume; plant; ds.
 XX Rhizobium sp.
 FE
 FT Key Location/Qualifiers
 FT CDS 417796..418671
 FT /tag= a
 FT /standard_name= "ORF K1"
 FT /product= "oligopeptide permease"
 FT /note= "homologous to the oppC gene"
 FT CDS 418673..419680
 FT /tag= b
 FT /standard_name= "ORF K2"
 FT /product= "oligopeptide permease"
 FT /note= "homologous to the oppD gene"
 FT CDS 419677..420738
 FT /tag= c
 FT /standard_name= "ORF K3"
 FT /product= "oligopeptide permease"
 FT /note= "homologous to the oppF gene"
 FT CDS 420774..422159
 FT /tag= d
 FT /standard_name= "ORF K4"
 FT /product= "encapsulation-like protein"
 FT /note= "homologous to the CapA gene"
 FT CDS 422628..424031
 FT /tag= e
 FT /standard_name= "ORF K5"
 FT /product= "aminotransferase-like protein"
 FT /note= "homologous to the BioA gene"
 FT CDS 424056..425594
 FT /tag= f

```

FT FT /standard_name= "ORF K6"
FT FT /product= "(semi)aldehyde dehydrogenase-like protein"
FT FT complement (426949..428028)
FT FT /*tag= g
FT FT /standard_name= "ORF K7"
FT FT /product= "transposase homologue"
FT FT /note= "homologous to the Tnp gene"
FT FT 428292..429623
FT FT /*tag= h
FT FT /standard_name= "ORF K8"
FT FT /product= "glutamate dehydrogenase-like protein"
FT FT /note= "homologous to the GLUD1 gene"
FT FT complement (430538..431284)
FT FT /*tag= i
FT FT /standard_name= "ORF K9"
FT FT /product= "transposase homologue"
FT FT complement (431296..432840)
FT FT /*tag= j
FT FT /standard_name= "ORF K10"
FT FT /product= "transposase homologue"
FT FT /note= "homologous to the Tnp gene"
FT FT complement (433880..434110)
FT FT /*tag= k
FT FT /standard_name= "ORF K11"
FT FT /product= "protein of unknown function"
FT FT /note= "homologous to the FixU gene"
FT FT complement (434107..434433)
FT FT /*tag= l
FT FT /standard_name= "ORF K12"
FT FT /product= "protein of unknown function"
FT FT complement (434517..434711)
FT FT /*tag= m
FT FT /standard_name= "ORF K13"
FT FT /product= "ferrodoxin/ferrodoxin-like protein"
FT FT /note= "homologous to the FdxN gene"
FT FT complement (434753..436234)
FT FT /*tag= n
FT FT /standard_name= "ORF K14"
FT FT /gene= "nifB"
FT FT /product= "protein involved in FeMo co-factor
FT FT biosynthesis"
FT FT complement (436460..438130)
FT FT /*tag= o
FT FT /standard_name= "ORF K15"
FT FT /gene= "nifA"
FT FT /product= "positive regulator of nif, fix and other
FT FT genes"
FT FT complement (438297..438590)
FT FT /*tag= p
FT FT /standard_name= "ORF K16"
FT FT /gene= "fixX"
FT FT /product= "protein required for nitrogenase activity"
FT FT complement (438605..439912)
FT FT /*tag= q
FT FT /standard_name= "ORF K17"
FT FT /gene= "fixC"
FT FT /product= "protein required for nitrogenase activity"
FT FT complement (439923..441032)
FT FT /*tag= r
FT FT /standard_name= "ORF K18"
FT FT /gene= "fixB"
FT FT /product= "protein required for nitrogenase activity"
FT FT complement (441042..441899)
FT FT /*tag= s
FT FT /standard_name= "ORF K19"
FT FT /gene= "fixA"
FT FT /product= "protein required for nitrogenase activity"
FT FT complement (442316..442636)
FT FT /*tag= t
FT FT /standard_name= "ORF K20"
FT FT /product= "protein of unknown function"
FT FT complement (443313..443879)
FT FT /*tag= u

```

```

FT FT /standard_name= "ORF K21"
FT FT /product= "protein of unknown function"
FT FT 444337..445029
FT FT /*tag= v
FT FT /standard_name= "ORF K22"
FT FT /product= "ferrodoxin-like protein"
FT FT /note= "homologous to the NifQ gene"
FT FT 445088..446602
FT FT /*tag= w
FT FT /standard_name= "ORF K23"
FT FT /gene= "dctA"
FT FT /product= "C4-dicarboxylate transport protein"
FT FT /note= "homologous to the DctAI gene"
FT FT 446599..447843
FT FT /*tag= x
FT FT /standard_name= "ORF L1"
FT FT /product= "cytochrome P450-like protein"
FT FT /note= "homologous to the CamC gene"
FT FT 447844..448500
FT FT /*tag= y
FT FT /standard_name= "ORF L2"
FT FT /product= "gamma-hexachlorocyclohexane dechlorinase-like
FT FT protein"
FT FT /note= "homologous to the LinA gene"
FT FT 448497..450203
FT FT /*tag= z
FT FT /standard_name= "ORF L3"
FT FT /product= "putative protein with degradative function"
FT FT 450341..451396
FT FT /*tag= aa
FT FT /standard_name= "ORF L4"
FT FT /product= "luciferase alpha-subunit-like protein"
FT FT /note= "homologous to the LuxA gene"
FT FT 452980..454494
FT FT /*tag= ab
FT FT /standard_name= "ORF L6"
FT FT /gene= "nifD"
FT FT /product= "alpha-subunit of FeMo protein of nitrogenase"
FT FT 454590..456131
FT FT /*tag= ac
FT FT /standard_name= "ORF L7"
FT FT /gene= "nifK"
FT FT /product= "beta-subunit of FeMo protein of nitrogenase"
FT FT 456187..457677
FT FT /*tag= ad
FT FT /standard_name= "ORF L8"
FT FT /product= "protein involved in FeMo co-factor
FT FT biosynthesis"
FT FT /note= "homologous to the NifB gene"
FT FT 457687..459096
FT FT /*tag= ae
FT FT /standard_name= "ORF L9"
FT FT /product= "protein involved in FeMo co-factor
FT FT biosynthesis"
FT FT /note= "homologous to the FixF gene"
FT FT 459093..459575
FT FT /*tag= af
FT FT /standard_name= "ORF L10"
FT FT /product= "protein of unknown function"
FT FT /note= "homologous to the NifX gene"
FT FT 459579..460067
FT FT /*tag= ag
FT FT /standard_name= "ORF L11"
FT FT /product= "protein of unknown function"
FT FT 460501..460920
FT FT /*tag= ah
FT FT /standard_name= "ORF L12"
FT FT /product= "protein similar to part of the Fe protein
FT FT of nitrogenase"
FT FT /note= "homologous to the NifH gene"
FT FT 461228..461545
FT FT /*tag= ai
FT FT /standard_name= "ORF L13"

```

|| |||:: || |
|| |||:: || ::
|| |||:: || |||
|| |||:: || |||

QY	4	IeuLysAsnLeuIleThrValValAlaPheSerIlePhe-----	Trip 17
DB	13	CTGACGACAGCCCTGTGTGTGTGGGACACGACCTGTTTGCCTTGGCAGTGTCTGGGTGGC 72	
QY	18	ValLeuLeuIleThrValAsnValTyrLeuPhe-----GlyAlaLysGlySerLeuSer 35	
DB	73	ATCCTGGCAGCCTATGTGTGACGGGTACCGAGTTTCATCCACACGGAAAGACATCACTGTCC 132	
QY	36	-----IleTyrGlyPheLeuLeuIleAlaTyrLeuLeuValLysMetSerLeuSerPhe 53	
DB	133	TTCGGCCTGTACGGGCCCATCTCGGGCTGCACCTGCTCATTCAGAGCCCTTTTGCCTTC 192	
QY	54	PheTyrLysProPheLysGlyArgAlaGlyGlnTyr-----	65
DB	193	CTGGAGCACCGGCGCATCGAGGTGCGCGGCAGCGCCCTGAAGCTGCCCTCCCGCGCGCG 252	
QY	66	---LysValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSerLeuLeuThr 84	
DB	253	GGCTCGGTGGCATGTGCATTCGCCGTACGAGGAGCCCTGACTACTTGGCGCAAGTGC 312	
QY	85	LeuLysSerValGlnGlnThrTyrProLeuAlaGluIleTyrValValAspAspGly 104	
DB	313	CTGCGCTCGGCCACGAGCGCATCTCCTCCCTGACCTCAAGGTGTCATGTGTGTGATGGC 372	
QY	105	SerAlaAspGlu-----	108
DB	373	AAACGCCAGGAGGACGCGCTACATGCTGGACATCTCCAGAGGTGCTGGCGGCGCACCGAG 432	
QY	108	-----	108
DB	433	CAGCGCGGCTCTTGTGTGGCGCAGCAACTCCATGAGGCGAGCGCAGGTGACACGAG 492	
QY	109	-----ThrGlyIleLysArgIleGluAspTyrValArgAspThrGlyAspLeu 124	
DB	493	GCCAGCCTGCAGGAGGCGATGACCGTGTCCGGATGTGTGCGG-----	537
QY	125	SerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArgHisAlaGlnAla 144	
DB	538	---GCCAGCATCTTCGTGCATCATCAGAGTGGGAGCGAAGCGCGAGTGCATGTAC 594	
QY	145	TrpAlaPheGluArg-----SerAspAlaAspValPheLeuThrValAspSerAspThr 162	
DB	595	ACGGCTTCACAGGCGCTCGCGCGATTCGGTGGACTACATCCAGGTGTGCGACTCTGCACACT 654	
QY	163	TyrIleTyrProAspAlaLeuGluIleLeuLysThrPhe---AsnAspProThrVal 181	
DB	655	GTGCTGGATCCAGCCCTGCACCATCGAGATCTTCGAGTCTCGGAGGAGGATCCCCAAGTA 714	
QY	182	PheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArgLeu 201	
DB	715	GGGGAGTCGGGGAGATGTCCAGATCCTCAACAGTACGACTCATGTGGATTCCTCTCTG 774	
QY	202	ThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThrGly 221	

PR 13-NOV-2001; 2001US-350666P.
 PR 29-NOV-2001; 2001US-334370P.
 PR 12-APR-2002; 2002US-372246P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.

PI Aziz N, Murray R;

DR WPI; 2003-093161/08.

DR P-PSDB; AB056686.

XX Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.

XX Claim 22; Page 401-402; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridises
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung
 CC cancer-associated polynucleotides and polypeptides are used for
 CC identifying a compound that modulates a lung cancer-associated
 CC polypeptide, for inhibiting proliferation of a lung cancer-associated
 CC cell to treat lung cancer in a patient and for treating a mammal having
 CC lung cancer by administering a modulatory compound identified. The
 CC methods are useful for treating lung cancer, such as small cell lung
 CC cancer, non-small cell lung cancer or other benign or precancerous
 CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
 CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial
 CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
 CC and polypeptides are useful for diagnostic purposes and as targets for
 CC screening for therapeutic compounds that modulate lung cancer, such as
 CC antibodies. Sequences ABX76124-ABX76474 represent lung cancer-associated
 CC polynucleotides of the invention.

XX SQ Sequence 1662 BP; 300 A; 501 C; 486 G; 375 T; 0 other;

Alignment Scores:

Pred. No.: 1-12e-36 Length: 1662
 Score: 447.50 Matches: 134
 Percent Similarity: 44.33% Conservative: 85
 Best Local Similarity: 27.13% Mismatches: 172
 Query Match: 20.86% Indels: 103
 DB: 25 Gaps: 16

US-09-469-200D-2 (1-417) x ABX76415 (1-1662)

QY 4 LeuLysAsnLeuIleThrValValAlaPheSerIlePhe-----Trp 17
 DB 13 CTGACGACAGCCCTGCGTGTGGTGCCACACAGCCTGTTGCCCTGGCAGTGTGGTGC 72
 QY 18 ValLeuIleThrValAsnValTyLeuPhe-----GlyAlaLysGlySerLeuSer 35
 DB 73 ATCCTGGCAGCCTATGTGAGCGGCTACCAAGTTATCCACGCGAAAGCACTACCTGTCC 132
 QY 36 -----IleTyGlyPheLeuLeuIleAlaTyLeuValLysMetSerLeuSerPhe 53
 DB 133 TTGGCGCTGTACGGCGCCATCCCTGGGCTGCACCTGCTATTACAGAGCCTTTTGCTTC 192
 QY 54 PheTyLysProPheLysGlyArgAlaGlyGlnTy----- 65
 DB 193 CTGGAGCACCGGCGCATCGAGCTGCGGCGCAGCCCTGAAGCTGCCCTCCCGCGGCGG 252
 QY 56 ---LysValAlaAlaIleProSerTyRasnGluAsnAlaGluSerLeuIleThr 84
 DB 253 GGCTCGGTGGCACTGTGCATTCGCCGTACAGAGGACCCCTGACTACTTGCACAAGTGC 312
 QY 85 LeuLysSerValGlnGlnThrTyRProLeuAlaGluIleTyRValValAspAspGly 104
 DB 313 CTCGCGTGGCGGACGCGATCTCCCTCCCTGACCTCAAGGTGTCATGGTGGTATGCG 372

QY 105 SerAlaAspGlu----- 108
 DB 373 AACCGCCAGGAGGAGCGCTACATGCTGGACATCTTCCACGAGGTGCTGGGGGACCGGAG 432
 QY 108 ----- 108
 DB 433 CAGCGCGGCTTCTTGTGTGGCGCAGCAACTTCCATGAGGCGGCGGCGGTGAGACGGAG 492
 QY 109 -----ThrGlyIleLysArgIleGluAspTyRValArgAspThrGlyAspLeu 124
 DB 493 GCCAGCCTGCGAGGAGGATGACCGGTGCGGATGTGTGGCG----- 537
 QY 125 SerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArgHisAlaGlnAla 144
 DB 538 ---GCCAGCACCTTCTCGTCATCATGCAAGTGGGAGGCAAGCGCGGCTCATGTAC 594
 QY 145 TrpAlaPheGluArg-----SerAspAlaAspValPheLeuThrValAspSerAspThr 162
 DB 595 ACGGCGCTTCAAGGCGCTCGGGATTCGGTGGACTACATCCAGGTGCGGACTCTGCACACT 654
 QY 163 TyrIleTyRProAsnAlaLeuGluGluLeuLeuLysThrPhe---AsnAspProThrVal 181
 DB 655 GTGTGTGATCCAGCCTCGACCATCGAGATGCTGAGTCTCGAGGAGGATCCCAAGTA 714
 QY 182 PheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArgLeu 201
 DB 715 GGGGAGTGGGGGAGATGTCAGATCCCTCAACAAAGTACGACTCATGGATTCTCTCTG 774
 QY 202 ThrAspIleArgTyRAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThrGly 221
 DB 775 AGCAGCGTGGGTACTGTGATGGCTTCAACGTGGAGGGGCGCTGCCAGTCTACTTTGGC 834
 QY 222 AsnIleLeuValCysSerGlyProLeuSerValTyRArgGluValValProAsn 241
 DB 835 TGTGTGCGAGTGTATAGTGGCGCCTTGGCATGTACCGCAACAGCCTCTCCACAGTTC 894
 QY 242 IleAspArgTyRileAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAspArg 261
 DB 895 CTGGAGGAGTGTGATCAGAGTCTTCCAGCAGCAAGTGCAGCTTCGGGGATGACCGG 954
 QY 262 CysLeuThrAsnTyRAlaThrAspLeuGly---LysThrValTyRGlinsThrAlaLys 280
 DB 955 CACCTCACCAACCGAGTCTCTGAGCTGGCTACCGAACTAAGTATACCGCGGCTCCCAAG 1014
 QY 281 CysIleThrAspValProAspLysMetSerThrTyRLeuLysGlnGlnAsnArgTyRAsn 300
 DB 1015 TGCCTCACAGAGACCCCACTAAGTACCTCGGTGCTCAACACGCAACCGCGCTGGAGC 1074
 QY 301 LysSerPhePheArgGluSerIle----- 308
 DB 1075 AAGTCTTACTCGGGAGTGGCTCTACAACCTCTCTGGTGTCCATTAAGCACCACTCTGG 1134
 QY 309 IleSerValLysIleMetAsnAsn-----ProPheValAlaLeuTyRThrIleLeu 326
 DB 1135 ATGACCTACAGTCAGTGGTGTTCACGGGTTCTTCCCTCTCTCTCTTCCATGCCAGGTTATA 1194
 QY 327 GluValSerMetPheMetMetLeuValTyRValValAspPhePheValGlyAsnVal 346
 DB 1195 CAGCTT-----TTCACCGGGCGCGCATC 1218
 QY 347 ArgGluPheAspTyRLeuArgValLeuAlaPheLeuValIleIlePheIleValAlaLeu 366
 DB 1219 -----TGG---AACATCTCCTTCTCTGTGACGGTGCAGCTGGTGGGCTT 1263
 QY 367 CysArgAsnIleHis---TyrMetLeuLysHisProLeuSerPheLeuLeuSerProPhe 385
 DB 1264 ATCAAGCCACCTACCTGCTTCTTCGGGCAATGCAGAGATGATCTTCATGCTCCCTC 1323
 QY 386 TyrGlyValLeuHisPheValGlnProLeuLysLeuTyRSerLeuPheThrIle 405
 DB 1324 TACTCCCTCTCTATATGTCAGCCCTTCTGCGGCGCAAGATCTTGGCATTTGCTACCATC 1383

QY 406 ArgAsnAlaAspTTPGlyThr-----ArgLysLysLeuLeu 417
 Db 1384 AACAAATCTGGCTGGGACACCTTGGCGGAAACCAATGTG 1425

RESULT 9

ABX76214
 ID ABX76214 standard; DNA; 4049 BP.

XX AC ABX76214;

XX DT 02-APR-2003 (first entry)

XX DE Lung cancer-associated polynucleotide #83.

XX KW Lung cancer-associated polynucleotide: gene; ds; cytostatic; emphysema;
 KW antiinflammatory; antilasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX OS Unidentified.

XX PN WO200286443-A2.

XX PD 31-OCT-2002.

XX PF 18-APR-2002; 2002WO-US12476.

XX PR 18-APR-2001; 2001US-284770P.
 PR 10-MAY-2001; 2001US-290492P.
 PR 09-NOV-2001; 2001US-339245P.
 PR 13-NOV-2001; 2001US-350666P.
 PR 29-NOV-2001; 2001US-334370P.
 PR 12-APR-2002; 2002US-372246P.

XX PA (EOSB-) EOS BIOTECHNOLOGY INC.

XX PI Aziz N, Murray R;

XX WPI; 2003-093161/08.
 DR P-PSDB; ABU56490.

XX PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer

XX PS Claim 22; Page 252-253; 453pp; English.

XX CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridises
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung
 CC cancer-associated polynucleotides and polypeptides are used for
 CC identifying a compound that modulates a lung cancer-associated
 CC polypeptide, for inhibiting proliferation of a lung cancer-associated
 CC cell to treat lung cancer in a patient and for treating a mammal having
 CC lung cancer by administering a modulatory compound identified. The
 CC methods are useful for treating lung cancer, such as small cell lung
 CC cancer, non-small cell lung cancer or other benign or precancerous
 CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
 CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial
 CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
 CC and polypeptides are useful for diagnostic purposes and as targets for
 CC screening for therapeutic compounds that modulate lung cancer, such as
 CC antibodies. Sequences ABX76214-ABX76474 represent lung cancer-associated
 CC polynucleotides of the invention.

XX SQ Sequence 4049 BP; 919 A; 1046 C; 1049 G; 1035 T; 0 other;

Alignment Scores:

Pred. No.: 3.92e-36 Length: 4049

Score:	447.50	Matches:	134
Percent Similarity:	44.33%	Conservative:	85
Best Local Similarity:	27.13%	Mismatches:	172
Query Match:	20.86%	Indels:	103
DB:	25	Gaps:	16
US-09-469-200D-2 (1-417) x ABX76214 (1-4049)			
QY	4	LeuLysAsnLeuLeuThrValValAlaPheSerIlePhe-----	-----TriP 17
Db	13	CTGACGACAGCCCTGCTGTGGTGGGCACACGCTGTTGCCCTGGCAGTGTGGGTGGC	72
QY	18	ValLeuLeuIleLeuValAsnValTyrLeuPhe-----GlyAlaLysGlySerLeuSer	35
Db	73	ATCCTGGCAGCTATGTGTACGGGCTACCCAGTTTATCCACACGAGAAAGCACTACTCTGC	132
QY	36	-----IleTyrGlyPheLeuLeuLeuAlaTyrLeuLeuValLysMetSerLeuSerPhe	53
Db	133	TTCCGCCCTGTACGGCGCCATCTCTGGCCCTGCACCTGCTCATTCAGAGCCTTTTTCCTTC	192
QY	54	PheTyrLysProPheLysGlyArgAlaGlyGlnTyr-----	65
Db	193	CTGGAGCACGGCGCATCGCACGTGCGCGCCCTGAAGCTGCCCTCCGCCGCGCG	252
QY	66	---LysValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSerLeuLeuThr	84
Db	253	GGCTCGGTGGCACTGTGCATTCGCGCATACAGAGGAGCCCTCACTACTTCCGAAGTGC	312
QY	85	LeuLysSerValGlnGlnInThrTyrProLeuAlaGluIleTyrValValAspGly	104
Db	313	CTGGCTCGGCCACGCGCATCTCTCCCTGACCTCAAGTGTCTATGGTGGTGGTGGC	372
QY	105	SerAlaAspGlu-----	108
Db	373	AACGCCAGGAGGACGCCCTACATCTTCCACGAGGTGCTGGCGGCGCACCGAG	432
QY	108	-----	108
Db	433	CAGCGCGGCTTCTTTGTGTGGCGCACAACTTCCATGAGCGAGGCGAGGTGAGACGGAG	492
QY	109	-----ThrGlyIleLysArgIleGluAspTyrValArgAspThrGlyAspLeu	124
Db	493	GCCAGCTTCAGAGGAGGCGATGGACCGTGTGGGATGTGTGGCGG-----	537
QY	125	SerSerAsnValIleValHisArgSerGluLysAsnGlnGlyArgHisAlaGlnAla	144
Db	538	---GCCAGCACCTTCTCTGTCATCATGCAGAAAGTGGGAGGCAAGCGCGGTCATGTAC	594
QY	145	TriAlaPheGluArg-----SerAspAlaAspValPheLeuThrValAspSerAspThr	162
Db	595	ACGGCCTTCAAGGCCCTCGCGGATTCGGTGACTACATCCAGGTGTGGCATCTTGACAT	654
QY	163	TyrIleTyrProAspAlaLeuGluLeuLeuLysThrPhe---AsnAspProThrVal	181
Db	655	GTGCTGGATCCACGCTCGCACCATCGAGATGCTCGAGTCTCTGGAGGAGATCCCCAAGA	714
QY	192	PheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnInThrAsnLeuLeuThrArgLeu	201
Db	715	GGGGAGTGGGGAGATGTCACAGATCTCTCAACAAGTACGACTCATGTGATTTCTCTCTG	774
QY	202	ThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThrGly	221
Db	775	AGCAGCGTGGGTACTGATGGCTTCAACGTGAGGGGCGCTGCCAGTCTCTACTTTGGC	834
QY	222	AsnIleLeuValCysSerglyProLeuSerValTyrArgGluValValProAsn	241
Db	835	TGTGTGCTAGCTATTAGTGGCCCTTGGCGCATGTACCGCAACAGCCCTCTCCACAGATC	894
QY	242	IleAspArgTyrIleAsnGlnInThrPheLeuGlyIleProValSerIleGlyAspArg	261
Db	895	CTGGAGGACTGGTACCATCAGATCTCTAGGCAGCAAGTGCAGCTTCGGGGATGACCGG	954
QY	262	CysLeuThrAsnTyrAlaThrAspLeuGly---LysThrValTyrGlnSerThrAlaLys	280

PT delivering drugs etc.
 XX
 PS
 Claim 11; Pages 75-78; 115pp; English.
 XX
 CC The present sequence encodes murine hyaluronan synthase-3 (HS-3),
 CC which can be used to alter hyaluronan, i.e. hyaluronic acid (HA),
 CC production, or for in vitro synthesis of HA.
 CC HA can be used to treat arthritis or perforated ear drums, protect
 CC eyes during surgery, deliver drugs, prevent post-operative scarring
 CC or adhesion formation, heal wounds and prevent restenosis following
 CC balloon angioplasty. Modulation of HA production in vivo may be
 CC useful in, e.g. Graves disease, mesothelioma, Wilms' tumour and
 CC oedema associated with inflammation of lung and kidney, all of
 CC which are associated with elevated levels of HA in tissues or
 CC serum.
 XX
 SQ Sequence 1665 BP; 340 A; 453 C; 451 G; 421 T; 0 other;
 Alignment Scores:
 Pred. No.: 6,81e-36 Length: 1665
 Score: 440.00 Matches: 134
 Percent Similarity: 44.44% Conservative: 86
 Best Local Similarity: 27.07% Mismatches: 171
 Query Match: 20.51% Indels: 104
 DB: 19 Gaps: 18
 US-09-469-2000-2 (1-417) x AAV18822 (1-1665)
 QY 4 LeuLysAsnLeuLeuIhrValValAlaPheSerIlePhe-----Tyr 17
 Db 13 CTGACTACAGCCCTCGGTGTGGTGGCACCAGTCTGTTGGCCCTGGTAGTCTGGGAGG 72
 QY 18 ValLeuLeuIleTyrValAsnValTyrLeuPhe-----GlyAlaLysGlySerLeuSer 35
 Db 73 ATCTGGCGGCGCTATGTGACAGCTACAGTTTATCCACAGAAAGACACTACCTGTCC 132
 QY 36 -----IleTyrGlyPheLeuLeuIleAlaTyrLeuValLysMetSerLeuSerPhe 53
 Db 133 TTTGGCCTCTACGGTGCATCTGGGTCTACATCTGCTCATCCAGAGCCCTGTTTGCCTTC 192
 QY 54 PheTyrLysProPheLysGlyArgAlaGlyGln----- 64
 Db 193 CTGGAGCACCCTCGAATGCGCAGGCGAGGCGCCCTCAAGCTGCACCTGCCAGAGG 252
 QY 65 ---TyrLysValAlaAlaIleIleProSerTyrAsnGluAspAlaGlnSerLeuLeuGlu 83
 Db 253 TCGCGTTCACTGGGACCTCTGCATCTGCTGCTTACCAGAGGACCCGAACTACCTGGCAG 312
 QY 84 ThrLeuLysSerValGlnGlnGlnThrTyrProLeuAlaGluIleTyrValValAsp 103
 Db 313 TGCCTTCGCTCAGTCAGCGCATTGCGCTTTCCAAACCTCAAGGTGGTCATGTTAGTGAT 372
 QY 104 GlySerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgAspThr----- 121
 Db 373 GGCATTCGCCAGAA-----GATACCTACATGTGGACATCTCCAT 414
 QY 122 -----GlyAspLeuSerSerAsnValIleValHisArgSer-----Glu 134
 Db 415 GAGGTGCTGGGTGGCCTCAGCAGACCTGGCTCTTTGTGTGGCGTAGCAATTCATGAG 474
 QY 135 LysAsnGlnGlyLysArgHisAla----- 142
 Db 475 GCGGGTGAAGGAGACAGAGGCCAGCCTCAGAGAAGCATGGAGCGTGGAGCTGTG 534
 QY 143 -----GlnAlaTrpAlaPheGluArg----- 149
 Db 535 GTGTGGCCAGCACCTTCTCATGTCATCATCAGAAAGTGGGGGGCAGCGTAGGTCTG 594
 QY 150 -----SerAspAlaAspValPheLeuThrValAspSerAsp 161
 Db 595 TACACTGCCTTCAGGCCCTTCAGGCACTCACTGGGACTACATCCAGGTGTGTGACTCTGAC 654
 QY 162 ThrTyrIleTyrProAspAlaLeuLeuGluLeuLeuLysThrPhe---AsnAspProThr 180

RESULT 10	
AAV18822	
ID	AAV18822 standard; cDNA; 1665 BP.
XX	
AC	AAV18822;
XX	
DT	11-JUN-1998 (first entry)
XX	
DE	Murine hyaluronan synthase-3 CDNA.
XX	
KW	Murine; hyaluronan synthase-3; HS-3; hyaluronan;
XX	hyaluronic acid; ss.
XX	
OS	Mus sp.
XX	
Key	Location/Qualifiers
FT	1..1665
CDS	/*tag= a
FT	/product= hyaluronan_synthase-3
FT	
FT	
FT	
PN	W09800551-A2.
XX	
PD	08-JAN-1998.
PD	
XX	
PD	03-JUL-1997; 97WO-US11761.
XX	
PD	
XX	
PR	05-MAR-1997; 97US-0812008.
PR	03-JUL-1996; 96US-0675499.
XX	
PA	(MAYO-) MAYO FOUNDATION.
XX	
PI	Augustine ML, McDonald JA, Spicer AP;
XX	
WPI	WPI; 1998-086976/08.
DR	P-PSDB; AAW50010.
DR	
XX	
PT	Nucleic acid encoding hyaluronan synthase-2 or -3 or their active
PT	fragments - used for treating arthritis, in wound repair, for


```

Db      415 GAGTCTGCTGGTGGCACTGAGCAAGCTGGCTTCTTGTGGCGTAGCAATTCATGAG 474
      |||      :|||      |||      |||      |||      |||
QY      135 LysAsnGlnGlyLysArgHisAla-----
      :|||      :|||      :|||      :|||      :|||      :|||
Db      475 GCGGGTGAAGGAGAGAGAGAGCGCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 534
      :|||      :|||      :|||      :|||      :|||      :|||
QY      143 -----GlnAlaTrpAlaPheGluArg----- 149
      :|||      :|||      :|||      :|||      :|||      :|||
Db      535 GTGTGGCCGACGACCTTCTCATGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAG 594
      :|||      :|||      :|||      :|||      :|||      :|||
QY      150 -----SerAspAlaAspValPheLeuThrValAspSerAsp 161
      :|||      :|||      :|||      :|||      :|||      :|||
Db      595 TACATGCGCTTCAAGGCCCTTGGCAACTCAGTGGACTACATCCAGGTGTGGACTCTGAC 654
      :|||      :|||      :|||      :|||      :|||      :|||
QY      162 ThrTyrIleTyrProAspAlaLeuGluGluLeuLysThrPhe---AsnAspProThr 180
      :|||      :|||      :|||      :|||      :|||      :|||
Db      655 ACTGTGCTGGACCCGCGCTGCACCATTCAGATTCGAGTCTTGGAGTCTTGAAGAGATCCCAA 714
      :|||      :|||      :|||      :|||      :|||      :|||
QY      181 ValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArg 200
      :|||      :|||      :|||      :|||      :|||      :|||
Db      715 GTAGGAGGTGTGGAGGAGATGTCCAAATCTCAACAGTATGATTCATGATCTCCTTC 774
      :|||      :|||      :|||      :|||      :|||      :|||
QY      201 LeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaGlnSerValThr 220
      :|||      :|||      :|||      :|||      :|||      :|||
Db      775 CTGAGCAGTGTGAGGTACTGGATGGCTTCAACGTGGAGGCGCTGCCAGTCTCTACTTT 834
      :|||      :|||      :|||      :|||      :|||      :|||
QY      221 GlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgArgGluValValPro 240
      :|||      :|||      :|||      :|||      :|||      :|||
Db      835 GCGTGTGTGCAATGATATGAGGCGCTTGGCGATGTACCGAACAGCGCTCTCAGCAG 894
      :|||      :|||      :|||      :|||      :|||      :|||
QY      241 AsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAsp 260
      :|||      :|||      :|||      :|||      :|||      :|||
Db      895 TTCTTGAGGATGGTACCATCAGAGTCTCTAGGACGAGCAAGTCAGCTTGGGAGTAT 954
      :|||      :|||      :|||      :|||      :|||      :|||
QY      261 ArgCysLeuThrAsnTyrAlaThrAspLeuGly---LysThrValTyrGlnSerThrAla 279
      :|||      :|||      :|||      :|||      :|||      :|||
Db      955 CGGCACCTTACCAACCGAGTCGTGAGTCTGTGCTACCGACTAAGTATACAGCACGCTCT 1014
      :|||      :|||      :|||      :|||      :|||      :|||
QY      280 LysCysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrp 299
      :|||      :|||      :|||      :|||      :|||      :|||
Db      1015 AAGTGGCTCAGAGACCCGCTAGTACCTTCGATGGCTCAATCAGCAACCCGCTGG 1074
      :|||      :|||      :|||      :|||      :|||      :|||
QY      300 AsnLysSerPhePheArgGluSerIle----- 308
      :|||      :|||      :|||      :|||      :|||      :|||
Db      1075 AGCAAGTCTTACTTTGGGAATGGCTCTACAAATTCCTGTGTTCCATAGCACCACTTC 1134
      :|||      :|||      :|||      :|||      :|||      :|||
QY      309 ---IleSerValLysLysIleMetAsnAsn-----ProPheValAlaLeuTrpThrIle 325
      :|||      :|||      :|||      :|||      :|||      :|||
Db      1135 TGGATGACCTATGAATCAGTGGTGCACAGGTTTCTCCCATTTCTCCTCATGTCTACAGTC 1194
      :|||      :|||      :|||      :|||      :|||      :|||
QY      326 LeuGluValSerMetPheMetMetLeuValTyrSerValValAspPhePheValGlyAsn 345
      :|||      :|||      :|||      :|||      :|||      :|||
Db      1195 ATCAAACTT-----TTCACCTGGCGCG 1218
      :|||      :|||      :|||      :|||      :|||      :|||
QY      346 ValArgGluPheAspTrpLeuArgValLeuAlaPheLeuValIlePheIleValAla 365
      :|||      :|||      :|||      :|||      :|||      :|||
Db      1219 ATC-----TGG---AACATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1263
      :|||      :|||      :|||      :|||      :|||      :|||
QY      366 LeuCysArgAsnIleHis---TyrMetLeuLysHisProLeuSerPheLeuLeuSerPro 384
      :|||      :|||      :|||      :|||      :|||      :|||
Db      1264 ATTATCAGGCTACTATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1323
      :|||      :|||      :|||      :|||      :|||      :|||
QY      385 PheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyrSerLeuPheThr 404
      :|||      :|||      :|||      :|||      :|||      :|||
Db      1324 CTCTACTCTCTCTCTATATATGCTGAGCTCTGCGCAGCAAGATCTTGTCTATTGCTACC 1383
      :|||      :|||      :|||      :|||      :|||      :|||
QY      405 IleArgAsnAlaAspTrpGlyThr-----ArgLysLysLeuLeu 417
      :|||      :|||      :|||      :|||      :|||      :|||
Db      1384 ATCAACAAGTCTGCTGGGCACTTCTGCGCAGGAACCACTTGTCT 1428
      :|||      :|||      :|||      :|||      :|||      :|||

```

RESULT 12
ABZ76736

ABZ76736 standard; cDNA; 1665 BP.
ABZ76736;
01-APR-2000 (first entry)
Mouse hyaluronan synthase 3 encoding cDNA SEQ ID NO:3.
Mouse; hyaluronan synthase; HAS; HAS3; enzyme; ophthalmological;
osteopathic; antiarthritic; gene therapy; angiogenesis inhibitor;
hyaluronic acid; dihydrazide; bioconjugate; dry eye syndrome; eye;
osteoarthritis; gene; ss.
Mus sp.
Key Location/Qualifiers
CDS 1..1665
/*tag= a
/product= "hyaluronan synthase 3 (HAS3)"
W02003006068-A1.
23-JAN-2003.
10-JUL-2001; 2001WO-US21785.
10-JUL-2001; 2001WO-US21785.
(CLEA-) CLEAR SOLUTIONS BIOTECH INC.
Dehaza P, Chen W;
WPI; 2003-221664/21.
P-PSDB; ABP96030.
Novel dihydrazide derivatized hyaluronic acid/nucleic acid bioconjugate
for treating dry eye syndrome, has derivatized hyaluronic acid
crosslinked to nucleic acid encoding protein with hyaluronan synthase
activity -
Claim 19; Page 57-58; 62pp; English.
The present invention describes a dihydrazide derivatized hyaluronic acid
(HA)/nucleic acid (NA) bioconjugate (I), with derivatized HA linked to NA
encoding protein with hyaluronan synthase (HAS) activity, where NA has
sequence with 79-85% sequence identity (SI) to a sequence of 1752, 1659
or 1665 nucleotides (see ABZ76734 to ABZ76736) or a sequence encoding
proteins with 66-90% SI to a sequence of 583, 552 or 554 amino acids (see
ABP96028 to ABP96030). (I) has ophthalmological, osteopathic and
antiarthritic activities, and can be used in gene therapy and as an
inhibitor of angiogenesis, and as an inducer of expression of (HA) in
human corneal epithelial cell. (I) is useful for transfecting a cell of
an eye with nucleic acid, where the nucleic acid comprises (ABZ76734 to
ABZ76736), by contacting the cell with (I) comprising the nucleic acid.
(I) is useful for treating dry eye syndrome in an individual. (I) is
useful in gene therapy applications for the treatment of a variety of
medical conditions including dry eye syndrome or other medical conditions
where an increase in the production of (HA) in the eye would be
therapeutic (e.g., osteoarthritis of the articular joints). (I) is also
useful for inhibiting angiogenesis for the treatment of macular
degeneration or genes related to lipid biosynthesis that helps to restore
the lipid component of the tear film, and as reagents for in vitro
transformation of any cell, preferably a eukaryotic cell, more preferably
a human eye cell. The present sequence encodes mouse HAS3 which is used
in the exemplification of the present invention.
Sequence 1665 BP; 340 A; 453 C; 451 G; 421 T; 0 other;
Alignment Scores:
Pred. No.: 6.81e-36 Length: 1665
Score: 440.00 Matches: 134
Percent Similarity: 44.44% Conservative: 86
Best Local Similarity: 27.07% Mismatches: 171

Query Match: 20.51% Indels: 104
DB: 25 Gaps: 18
US-09-469-200D-2 (1-417) x ABZ76736 (1-1665)

QY 4 LeuLysAsnLeuIleThrValValAlaPheSerIlePhe-----TTP 17
DB 13 CTGACACAGCCCTGCGTGTGGGACACCACTGTTGCGCTGAGTGGGAGGC 72
QY 18 ValLeuLeuIleThrValValAlaPheSerIlePhe-----GlyAlaIleSerLeuSer 35
DB 73 ATCTGGGGCTGAGTACAGAGCTACAGGTTATCCACACAGAAAAGCACTACCTGTCC 132
QY 36 -----IleTyrGlyPheLeuLeuIleAlaTyrLeuValValAlaPheSerIlePhe 53
DB 133 TTGGGCTCTACGCTGCGTGTGGGACACCACTGCTCATCCAGAGCGTGTGGCTTC 192
QY 54 PheTyrLysProPheLysGlyArgAlaGlyCln----- 64
DB 193 CTGGAGCACCTCGAATGCGCAGGCGGCGCCCTCAAGCTGCACTGCTCCAGAGG 252
QY 65 ---TyrLysValAlaAlaIleProSerTyrAsnGluAspAlaGluSerLeuLeuGlu 83
DB 253 TCGCTTCAGTGGGCACTGCTGCTGCTTACAGAGACCCCGGATACCTTCCGCAAG 312
QY 84 ThrLeuLysSerValGlnGlnThrTyrProLeuAlaGluIleTyrValValAlaPhe 103
DB 313 TGCTTCGCTCAGTCAGGCGCATTCCTTCCAAACCTCAAGGTGCTGATGGAT 372
QY 104 GlySerAlaAspGluThrGlyLysArgIleGluAspTyrValArgAspThr----- 121
DB 373 GGCATCCCGAGNA-----GATACCTACATGTTGGACATCTTCCAT 414
QY 122 -----GlyAspLeuSerSerAsnValIleValHisArgSer-----Glu 134
DB 415 GAGGTGCTGGTGGCACTGACGAGCTGGCTTCTTGTGGCTGACATTTCCATGAG 474
QY 135 LysAsnGlnGlyArgHisAla----- 142
DB 475 GCGGTGAAGCAGACAGACAGCCAGCGCTGAGAGGCGTGGCGAGCTGTG 534
QY 143 -----GlnAlaTyrAlaPheGluArg----- 149
DB 535 GTGTGGCCAGCACCTTCTCATGCAATCATGCAAGTGGGGGCGCAAGCTGAGTCAIG 594
QY 150 -----SerAspAlaAspValPheLeuIleThrValAlaPheSerAsp 161
DB 595 TACACTGCTTCAAGGCCCTTGGCACTGACGTACATCCAGGTGTGCTGCTGAC 654
QY 162 ThrTyrIleTyrProAspAlaLeuGluLeuLysThrPhe---AsnAspProThr 180
DB 655 ACTGTCTGGAGCCCGCTGACCATTCAGATGCTTCCAGTCTTGGAGAGATCCCAA 714
QY 181 ValPheAlaAlaThrGlyHisAsnValArgAsnArgGlnThrAsnLeuLeuThrArg 200
DB 715 GTAGAGGTGTGGAGGAGATGCTCAATCTCCACAGATGATGATTCAGTGAATCTCTC 774
QY 201 LeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThr 220
DB 775 CTGAGAGTGTGAGGTGAGTGGATGCTTCAAGCTGGAGCGGCTGCCAGTCTTACTTT 834
QY 221 GlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgGluValValPro 240
DB 835 GCTGTGTGCAATGATATGAGTGGGCTTGGGCGATGTACCCCAACAGCCCTTCCAGAG 894
QY 241 AsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAsp 260
DB 895 TTCTTGGAGGATGTGATACCATCAGAGTCTTACGACAGCAGTGGAGTGGGATGAT 954
QY 261 ArgCysLeuThrAsnTyrAlaThrAspLeuGly---LysThrValTyrGlnSerAla 279
DB 955 CGGCACCTTACCAACCGAGTCTGAGTCTGGCTACCGGACTAAGTATACAGCAGCTCT 1014

QY 280 LysCysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTyr 299
DB 1015 AAGTGCCTCAGACAGCCGCCCACTAGCTACCTTCGATGGCTCAATCAGCAACCCGCTGG 1074
QY 300 AsnLysSerPhePheArgLysIlePhe----- 308
DB 1075 AGCAGCTTACTTTCGGGAATGGCTTCAATTCCTCTGTGGTTCATTAAGCACCACCTC 1134
QY 309 ---IleSerValLysIleMetAsnAsn-----ProPheValAlaLeuTyrThrIle 325
DB 1135 TGGATGACCTATCAATCAGTGGTCCAGAGTTCCTTCCCATCTTCTCATGCTACAGTC 1194
QY 326 LeuGluValSerMetPheMetLeuValTyrSerValValAspPhePheValGlyAsn 345
DB 1195 ATACAACTT-----TTTACCGTGGCGG 1218
QY 346 ValArgGluPheAspTyrLeuArgValLeuAlaPheLeuValIleIlePheIleValAla 365
DB 1219 ATC-----TGG---AACATTCCTCTTCTGTGTACAGAGCTGGTGGG 1263
QY 366 LeuCysArgAsnIleHis---TyrMetLeuLysHisProLeuSerPheLeuLeuSerPro 384
DB 1264 ATTATCAAGGCTACCTATGCTGCTTCTGAGGCAATCGACAGATGATCTTCAATGTC 1323
QY 385 PheTyrGlyValIleHisLeuPheValLeuGlnProLeuLysLeuTyrSerLeuPheThr 404
DB 1324 CTCTACTCTCTCTATATGTCCAGCTTGTGCCCAAGATCTTGTCTATTGCTACC 1383
QY 405 IleArgAsnAlaAspTyrGlyThr-----ArgLysLysLeuLeu 417
DB 1384 ATCAACAAGCTCTGGTGGGCACTTCTGCGAGAAACCATTTGTC 1428

RESULT 13
ABLA1013/c
ID ABLA1013 standard; cDNA; 1653 BP.
AC ABLA1013;
XX AC
XX AC
DT 30-JUL-2002 (first entry)
XX Murine hyaluronan synthetase 2 (HAS2) cDNA fragment.
DE Hyaluronan synthetase 2, HAS2; tumour; gene therapy; mouse; gene; ss.
KW Mus sp.
XX Mus sp.
OS JP2002065280-A.
XX JP2002065280-A.
PD 05-MAR-2002.
XX 05-MAR-2002.
PF 01-SEP-2000; 2000JP-0266260.
XX 01-SEP-2000; 2000JP-0266260.
PR 01-SEP-2000; 2000JP-0266260.
XX (SEK) SEIKAGAKU KOGYO CO LTD.
XX WPI; 2002-397661/43.
XX A DNA vector for inhibiting the formation of a malignant tumor -
XX Claim 1; Page 6; 9pp; Japanese.
XX The invention relates to a DNA vector containing a 20 nucleotide sequence
XX of the 3'-end of a mouse hyaluronan synthetase 2 (HAS2) cDNA sequence,
XX or DNA hybridisable with a polynucleotide complementary to it under
XX stringent conditions, and which can inhibit the translation of a
XX hyaluronan synthetase 2 gene in which the transcription initiating point
XX is recombined 5' upstream of the DNA. The vector is used for inhibiting
XX the formation of a malignant tumour. The present sequence represents a
XX mouse HAS2 cDNA fragment.
XX Sequence 1653 BP; 471 A; 387 C; 353 G; 442 T; 0 other;

Alignment Scores:

Pred. No.: 8-57e-36 Length: 1653
 Score: 439.00 Matches: 134
 Percent Similarity: 46.65% Conservativeness: 82
 Best local Similarity: 28-94% Mismatches: 183
 Query Match: 20-47% Indels: 65
 DB: 24 Gaps: 14

US-09-469-200D-2 (1-417) x ABL41013 (1-1653)

QY 7 LeuLeuThrValValAlaPheSerIlePheTrpValLeuLeuIleTyrValAsnValTyr 26
 Db 1590 TCCTCCGGAATCACAGCTGCTTATATATTGGCTACCAAGTTTATCCAAACAGATAATAC 1531
 QY 27 LeuPheGlyAlaLysSerLeuSerIleTyrGlyPheLeuLeuIleAlaTyrLeuLeu 46
 Db 1530 TACTTC-----TCATTGGACTGTAGGTGCTTTTATCCCTCGCATCTCATC 1483
 QY 47 Val-LysMetSerLeuSerPhePhe-----TyrLysProPheLysGlyAr 61
 Db 1482 ATCCAAAGCCTCTTTCCTTTTGGAAACACCGGAAATGAAGAAGTCCCTTGAAACCCCG 1423
 QY 61 AlaGlyGlnTyrLys---ValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSe 80
 Db 1422 ATTAATG-ACAAAACGGTAGCATCTGTCATCGCTCGTCCGACCAAGAGACCCCTGACTA 1364
 QY 80 rLeuLeuGluThrLeuLysSerValGlnGlnGlnThrTyrProLeuAlaGluIleTyrVa 100
 Db 1363 CTTACGGAATGTTGCAATCTGTGAAGAGCTGACCTACCTCCCTGGGATTAAGTCGTGAT 1304
 QY 100 lValAspAspGlySerAlaAsp----- 107
 Db 1303 GGTATCATGATGGAATCATGACAGCAGCAGCCTTTTACATGATGACATATTCAGCGAAGTAT 1244
 QY 108 -----GluThrGlyIleLy 112
 Db 1243 GGCAGGCGCAATCGCCACGCTACATCTGGGAAGAACACTTTCATGAAGAGGACCTCG 1184
 QY 112 sArgIleGluAspTyrValArgAspThrGlyAsp-----LeuSerSerAs 127
 Db 1183 TGAGACAGAGAGATCCCATAAAGAAATTCACACATGTCACCCAAATGGTCTTGCTTAA 1124
 QY 127 nValIleValHisArgSerGluLysAsnGlnGlyLysArgHisAlaGlnAlaTrpAlaph 147
 Db 1123 CAAGAATATTGTCATCATGCAAAATGGGGTGGGAAGAGAGAGATCATGTACACAGCCTT 1064
 QY 147 eGluArg-----SerAspAlaAspValPheLeuThrValAspSerAspThrTyrIleTy 165
 Db 1063 CAGAGCAGCTGGGCGCAAGCGTGGATTATGTACAGGTGTGTGACTCAGACTACTGTCTGA 1004
 QY 165 rProAspAlaLeuGluGluLeuLeuLysThrPhe---AsnAspProThrValPheAlaAl 184
 Db 1003 CCCTGCTCATCTGTTGGAGATGTTGAAGGCTTTAGAGGAAGACCCCTATGTTGGAGGTG 944
 QY 184 aThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArgLeuAspIl 204
 Db 943 TGGAGGAGATGCCAGATTTAAACAGATATGATTCCTGGATCTCTCTCCAGCAGCGT 884
 QY 204 eArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThrGlyAsnIleLe 224
 Db 883 GAGATACTGATGGCTTTTAAATATAGAAAGGCGCTGCCAGCTCTTATTTGGCTGTGCCA 824
 QY 224 uValCysSerGlyProLeuSerValTyrArgArgGluValValProAsnIleAspAr 244
 Db 823 GTGCATAAAGCGCTCTCGGAATGTACAGAAACTCTCTGTGCTGATGATTTGTGAAGA 764
 QY 244 gTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAspArgCysLeuTh 264
 Db 763 CTGTATCATCAGGAATTCATGGGTAACCAATGCGATTTTGGTGACGACGACCCCTTAC 704
 QY 264 rAsnTyrAlaThrAspLeuGly---LysThrValTyrGlnSerThrAlaLysCysIleTh 283
 Db 703 CAACAGGCTGTTGAGTCTGGGCTATGCAACTAAATACACGGCTCGTCCAGTGCCTTAC 644

QY 283 rAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrpAsnLysSerPh 303
 Db 643 TGAACACTCCCATAGAATATCTGAGATGGCTGAACACAGACCCGTTGGAGCAAGTCTTA 584
 QY 303 ePheArgGluSerIleIleSerValLysLysIleMetAsnAsnProPheValAlaLeuTr 323
 Db 583 CTTCGACAGCTGGCTGTACAAAGTGGTTCACAAAGCAT-----CACCTGTG 533
 QY 323 pThrIleLeuGluValSerMet-----PheMetMetLeuValTyrSerVa 338
 Db 532 GATGACCATATCAAGCTGTTATCACTGGATTCTTTCTTCTTCTTCATTCGACAGTCAT 473
 QY 338 lValAspPhePheValGlyAsnValArgGluPheAspTrpLeuArgValLeuAlaPheLe 358
 Db 472 CCAGCTCTTCACAGGGGTAATC-----TGG---AACATCTCTCTCTCTCT 428
 QY 358 uValIleIlePheIleValAlaLeuCysArg---AsnIleHisTyrMetLeuLysHisPr 377
 Db 427 GTTAACTGTCCAGCTAGTGGTCTCTCACTCAAGTCATCTTTGCCAGCTGCCTTAGAGGAAA 368
 QY 377 oLeuSerPheLeuLeuSerProPheTyrGlyValLeuHisLeuPheValLeuGlnProLe 397
 Db 367 TATCCTCATGTTATTCATGCTCTCTATTCAGTGTATATACATGTCAAAGTCTACTTCTGC 308
 QY 397 uLysLeuTyrSerLeuPheThrIleArgAsnAlaAspTrpGlyThr-----ArgLysLy 415
 Db 307 CAAAGTGTTCAAATTCGAACCATTAACAAGCTGGGTGGGCGACATCTGGAGGAAGAC 248
 QY 415 sLeuLeu 417
 Db 247 CATTGTT 241
 RESULT 14
 ABZ76735
 ID ABZ76735 standard; cDNA; 1659 BP.
 AC XX
 ABZ76735;
 DT XX
 01-APR-2000 (first entry)
 DE Mouse hyaluronan synthase 2 encoding cDNA SEQ ID NO:2.
 KW Mouse; hyaluronan synthase; HAS; HAS2; enzyme; ophthalmological;
 KW osteopathic; antiarthritic; gene therapy; angiogenesis inhibitor;
 KW hyaluronic acid; dihydrazide; bioconjugate; dry eye syndrome; eye;
 KW osteoarthritis; gene; ss.
 XX OS
 Mus sp.
 XX Key Location/Qualifiers
 FH 1..1659
 FT CDS /*tag= a
 FT /product= "hyaluronan synthase 2 (HAS2)"
 XX WO2003006068-A1.
 XX 23-JAN-2003.
 XX 10-JUL-2001; 2001WO-US21785.
 XX 10-JUL-2001; 2001WO-US21785.
 PR (CLEA-) CLEAR SOLUTIONS BIOTECH INC.
 XX Dehazya P, Chen W;
 DR WPI; 2003-221664/21.
 DR P-ESDB; ABP96029.
 XX Novel dihydrazide derivatized hyaluronic acid/nucleic acid bioconjugate
 PT for treating dry eye syndrome, has derivatized hyaluronic acid
 PT crosslinked to nucleic acid encoding protein with hyaluronan synthase

PT activity -

Claim 19; Page 56-57; 62pp; English.

CC The present invention describes a dihydrazone derivatised hyaluronic acid (HA)/nucleic acid (NA) bioconjugate (I), with derivatised HA linked to NA encoding protein with hyaluronan synthase (HAS) activity, where NA has sequence with 79-85% sequence identity (SI) to a sequence of 1752, 1659 or 1665 nucleotides (see AB276734 to AB276736) or a sequence encoding proteins with 66-90% SI to a sequence of 583, 552 or 554 amino acids (see AB276734 to AB276736). (I) has ophthalmological, osteopathic and antiarthritic activities, and can be used in gene therapy and as an inhibitor of angiogenesis, and as an inducer of expression of (HA) in human corneal epithelial cell. (I) is useful for transfecting a cell of an eye with nucleic acid, where the nucleic acid comprises (AB276734 to AB276736), by contacting the cell with (I) comprising the nucleic acid. (I) is useful for treating dry eye syndrome in an individual. (I) is useful in gene therapy applications for the treatment of a variety of medical conditions including dry eye syndrome or other medical conditions where an increase in the production of (HA) in the eye would be therapeutic (e.g., osteoarthritis of the articular joints). (I) is also useful for inhibiting angiogenesis for the treatment of macular degeneration or genes related to lipid biosynthesis that helps to restore the lipid component of the tear film, and as reagents for in vitro transformation of any cell, preferably a eukaryotic cell, more preferably a human eye cell. The present sequence encodes mouse HAS2 which is used in the exemplification of the present invention.

XX SQ Sequence 1659 BP; 445 A; 353 C; 388 G; 473 T; 0 other;

Alignment Scores:

Pred. No.: 8,62e-36 Length: 1659
Score: 439.00 Matches: 134
Percent Similarity: 46.65% Conservative: 82
Best Local Similarity: 28.94% Mismatches: 183
Query Match: 20.47% Indels: 65
DB: 25 Gaps: 14

US-09-469-200D-2 (1-417) x AB276735 (1-1659)

QY 7 LeuileThrValAlaPheSerIlePheTrpValLeuLeuIleTyrValAsnValTyr 26
Db 67 CTCCTCGGAATCACAGCTGCTTATATTTGGCTACCATTTATCCAAACAGATAATATAC 126
QY 27 LeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeuLeuIleAlaTyrLeuLeu 46
Db 127 TACTTC-----TCATTTGGAGTGTAGGTCCTTTTATAGCTCGCATCAATC 174
QY 47 Val-LysMetSerLeuSerPhePhe-----TyrLysProPheLysGlyAr 61
Db 175 ATCCAAAGCCCTTTGGCTTTTGGACACCGGAATGAAGAGTCCCTTGAACCCCG 234
QY 61 gAlaGlyInTyrLys---ValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSe 80
Db 235 ATTAAATTG-AACAAACGGTAGCAGCTCTGCATCGCTGGTACCAAGAGACCTGACTA 293
QY 80 rLeuLeuGluThrLeuLysSerValGlnGlnInThrTyrProLeuAlaGluLeyrVa 100
Db 294 CTTACGGAAATGTTGCAATCTGTGAAGAGCTGACCTACCTGGGATTAAAGTCGTGAT 353
QY 100 lValAspAspGlySerAlaasp----- 107
Db 354 GGTCAATGATGGAACTCAGACAGACGACCTTACATGATGGACATATTCAGCAAGTAT 413
QY 108 -----GluThrGlyIlely 112
Db 414 TGGCAGGCAAAATCGGCCACGTACATCTGGAAGAACAACTTTTCATGAAAGGACCTGG 473
QY 112 sArgIleGluAspTyrValArgAspThrGlyAsp-----LeuSerSerAs 127
Db 474 TGACAGAGAGAGTCCCATAAAGAAAGTTCACAACTATGTCACCAATTTGCTGCTAA 533
QY 127 nValIleValHisArgSerGluLysAsnGlnGlyLysArgHisAlaGlnAlaTrpAlaph 147

Db 534 CAAAAGTATTTCATCATCGCAAAATGGGTGAAAGAGAAAGTCAATGTACACAGCCTT 593
QY 147 ecluArg-----SerAspAlaAspValPheLeuThrValAspSerAspThrTyrIleTy 165
Db 594 CAGAGCACTGGGCGAAGCGTGGATTATGTACAGGTGTGTACTCAGATACTTACGTGTGA 653
QY 165 rProAspAlaLeuGluGluLeuLeuLysThrPhe---AsnAspProThrValPheAlaAl 184
Db 654 CCCTCGCTCATCTGTGGAGTGTGAAGTCTTAGAGGAGACCCCTATGGTTGGAGGTGT 713
QY 184 aThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArgLeuThrAspII 204
Db 714 TGGAGGAGATGTCAGCAATTTAAACAAAGTATGATTCCTGGATCTCTCTCAGCAGCGT 773
QY 204 eArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThrGlyAsnIleLe 224
Db 774 GAGATACCTGGATGCTTTTAAATATAGAAAGGCGCTGCCAGTCTTATTTGGGTGTGCCA 833
QY 224 uValCysSerGlyProLeuSerValTyrArgGluValValProAsnIleAspAr 244
Db 834 GTGCATAGCGGTCTCTGGGATGTACAGAACTCTCTGCTGCATGAATTTGTGAAGA 893
QY 244 gTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAspGlyCysLeuTh 264
Db 894 CTGGTACATCAGGAATTCATGGGTAAACCAATGCAGTTTGGTGACGACGACGACCTTAC 953
QY 264 rAsnTyrAlaThrAspLeuGly---LysThrValTyrGlnSerThrAlaLysCysIleTh 283
Db 954 CAACAGGGTGTGGTCTGGCTATGCAATTAATACACGGCTCGCTCAAGTGCCTTAC 1013
QY 283 rAspValProAspLysMetSerThrTyrLeuLysGlnGluAsnArgTyrAsnLysSerPh 303
Db 1014 TGAACCTCCCATAGATATCTGAGATGGCTGAACCGACGACGCCGATGAGCAAGTCTTA 1073
QY 303 ePheArgGluSerIleIleSerValLysLysIleMetAsnAsnProPheValAlaLeuTr 323
Db 1074 CTTCCGAGAGTGGCTGTACAATGCCATGTGGTTTTCACAGCAT-----CACCTGTG 1124
QY 323 pThrIleLeuGluValSerMet-----PheMetMetLeuValTyrSerVa 338
Db 1125 GATGACCTATGAGAGTGTATTCATCGATGATCTTCTCTTCTTCTCATTTGCCACAGTCA 1184
QY 338 lValAspPhePheValGlyAsnValArgGluPheAspTrpLeuArgValLeuAlaPheLe 358
Db 1185 CCACCTCTCTACAGGGTAAATC-----TGG---AACCTCTCTCTCTCTCT 1229
QY 358 uValIleIlePheIleValAlaLeuCysArg---AsnIleHisTyrMetLeuLysHisPr 377
Db 1230 GTTAACCTGCCAGCTAGTGGTCTCATCAAGTCACTTTTGGCAGCTGCTTAGAGGAA 1289
QY 377 oLeuSerPheLeuLeuSerProPheTyrGlyValLeuHisLeuPheValLeuGlnProLe 397
Db 1290 TATCGTATGATTCATGCTCTGTATTCAGTGTATACATCTCAAGTCTACTTCTCTG 1349
QY 397 uLysLeuTyrSerLeuPheThrIleArgAsnAlaAspTrpGlyThr-----ArgLysLy 415
Db 1350 CAAGATGTTGCAATTGCAACCATAAACAAGCTGGGTGGGCGACATCTGGAAGAGAGAC 1409
QY 415 sLeuLeu 417
Db 1410 CATTTGT 1416
RESULT 15
AAV18821
ID AAV18821 standard; cDNA; 2947 BP.
XX
AC AAV18821;
XX
DT 11-JUN-1998 (first entry)
XX
DE Murine hyaluronan synthase-2 cDNA.
XX

Db 1857 CAGATGTTTGCATTCGAACCATAAACAAGCTGGTGGGGCACATCTGGAAGGAGAC 1916
Qy 415 sLeuLeu 417
Db 1917 CATTGTT 1923

Search completed: October 2, 2003, 12:56:38
Job time : 698 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2003, 12:44:58 ; Search time 82 Seconds

(without alignments)

2244.595 Million cell updates/sec

Title: US-09-469-200D-2

Perfect score: 2145

Sequence: 1 MRLKNIITVAPISIFVLL.....KLYSLTIRNADWCTRKLL 417

Scoring table:

BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Fgapop 10.0 , Fgapext 0.5

Delop 6.0 , Delext 7.0

Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09469200/runat_02102003_090348_22797/app_query.fasta.1.583
-DB=Issued_Patents_NA -QWAF=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09469200@cgn.1.1.56 @runat_02102003_090348_22797 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA:*
1: /cgn2.6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2.6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2.6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2.6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2.6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2.6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1586.5	74.0	1512	2 US-08-270-581-1
2	1586.5	74.0	1512	4 US-09-146-893-1
3	463	21.6	536165	4 US-09-214-808-1
4	439	20.5	2947	4 US-08-675-499A-1
5	368.5	17.2	2117	3 US-09-155-768-1
6	366.5	17.1	2116	4 US-08-635-552A-1
7	242.5	11.3	4951	2 US-08-867-030B-5
8	242.5	11.3	4951	5 PCT-US95-06119-5
9	239.5	11.2	1326	4 US-09-107-532A-2775
10	239	11.1	1260	4 US-09-328-352-2963
11	234.5	10.9	1257	4 US-09-134-001C-1214
12	229	10.7	1263	4 US-09-328-352-1781

c	13	175	8.2	19390	4	US-08-961-527-86	Sequence 86, Appl
	14	166	7.7	4066	4	US-09-499-203-16	Sequence 16, Appl
	15	166	7.7	9321	4	US-09-499-203-1	Sequence 1, Appl
c	16	163	7.6	2673	4	US-09-252-991A-2993	Sequence 2993, Ap
	17	163	7.6	2715	4	US-09-252-991A-2686	Sequence 2686, Ap
	18	163	7.5	2814	4	US-09-252-991A-2875	Sequence 2875, Ap
	19	149	6.9	3084	1	US-08-551-437-1	Sequence 1, Appl
	20	149	6.9	3084	3	US-09-004-225-1	Sequence 1, Appl
	21	149	6.9	3084	3	US-09-084-346-1	Sequence 1, Appl
	22	149	6.9	3084	3	US-09-104-704-1	Sequence 1, Appl
	23	146.5	6.8	910	3	US-08-961-083-201	Sequence 201, App
	24	146.5	6.8	910	4	US-09-536-784-201	Sequence 201, App
	25	146	6.8	1311	4	US-09-328-352-290	Sequence 290, App
	26	144.5	6.7	1654976	4	US-08-916-421B-1	Sequence 1, Appl
	27	144	6.7	1830121	4	US-09-557-884-1	Sequence 1, Appl
	28	144	6.7	1830121	4	US-09-643-990A-1	Sequence 1, Appl
	29	143.5	6.7	912	4	US-09-252-991A-12584	Sequence 12584, A
c	30	140	6.5	8876	4	US-08-961-527-170	Sequence 170, App
	31	139.5	6.5	2166	4	US-09-107-532A-3235	Sequence 3235, Ap
	32	138	6.4	235	4	US-08-675-499A-25	Sequence 25, Appl
	33	137.5	6.4	522	4	US-08-858-207A-232	Sequence 232, App
c	34	137.5	6.4	1128	4	US-09-252-991A-1077	Sequence 1077, Ap
	35	137.5	6.4	1191	4	US-09-252-991A-1025	Sequence 1025, Ap
	36	137.5	6.4	3114	4	US-09-252-991A-977	Sequence 977, App
	37	135	6.3	235	4	US-08-675-499A-26	Sequence 26, Appl
	38	131	6.1	235	4	US-08-675-499A-23	Sequence 23, Appl
	39	128	6.0	506	4	US-09-252-991A-16525	Sequence 16525, A
	40	128	6.0	3060	4	US-09-115-150-1	Sequence 1, Appl
	41	127	5.9	951	4	US-09-107-532A-1037	Sequence 1037, Ap
	42	127	5.9	14602	1	US-08-597-236-1	Sequence 1, Appl
	43	127	5.9	14602	1	US-08-746-682A-1	Sequence 1, Appl
c	44	124.5	5.8	14187	4	US-09-453-702B-121	Sequence 121, App
	45	124	5.8	1026	4	US-09-328-352-3052	Sequence 3052, Ap

ALIGNMENTS

RESULT 1
US-08-270-581-1
; Sequence 1, Application US/08270581
; Patent No. 5856168
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H.
; APPLICANT: DeAngelis, Paul L.
; APPLICANT: Papaconstantinou, John
; TITLE OF INVENTION: Hyaluronate Synthase Gene and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/270,581
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTSG:161\PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-787-1400
; TELEFAX: 713-789-2679
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1512 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 193..1449
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1488..1510
 ;
 ; US-08-270-581-1

Alignment Scores:

Pred. No.: 4.66e-187 Length: 1512
 Score: 1586.50 Matches: 298
 Percent Similarity: 85.19% Conservatives: 53
 Best Local Similarity: 72.33% Mismatches: 60
 Query Match: 73.96% Indels: 1
 DB: 2 Gaps: 1

US-09-469-200D-2 (1-417) x US-08-270-581-1 (1-1512)

QY 5 LysAsnLeuLeuThrValValAlaPheSerIlePheTrpValLeuLeuLeuTyrValAsn 24
 DB 205 AAAAAAAGCTTAATGCTTTTATCCCTTTATTTTTCATATCTATCTTGATTTATCTAAAT 264
 QY 25 ValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeuLeuLeuLeuTyr 44
 DB 265 AGTATCTATTTGGACA---TCAACTGTAGGAATTTATGGAGTAATATTATTAACCTAT 321
 QY 45 LeuLeuValLysMetSerLeuSerPhePheTyrLysProPheLysGlyArgAlaGlyGln 64
 DB 322 CTAGTTATCAAACTTGATATCTTCTTTTATGAGCCATTTAAAGGAATCCCATGAC 381
 QY 65 TyrLysValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSerLeuLeuGluThr 84
 DB 382 TATAAAGTGTCTGCTTAATCTCTTATTAATGAGATGCGGAGTCATATTAGAAACA 441
 QY 85 LeuLysSerValGlnGlnThrTyrProLeuAlaGluIleTyrValValAspAspGly 104
 DB 442 CTTAAAGTGTGTACACAGACCTATCCGTTATCAGAAATTTATTTGATGATGGG 501
 QY 105 SerAlaAspGluThrGlyLysArgIleGluAspTyrValArgAspThrGlyAspLeu 124
 DB 502 AGTTCAACACAGATGCAATCAATTAATGAGAGTATGTAATAGAGAAGTGGATATT 561
 QY 125 SerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArgHisAlaGlnAla 144
 DB 562 TGTCAAAAGCTTATCTTCACGTTCCCTGTCATTAAGGAACCCATGCTCAAGCG 621
 QY 145 TrpAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSerAspThrTyrIle 164
 DB 622 TGGCATTTGAAAGATCTGACGCTGACGTTTTTAAACGCTAGACTCAGATCTTATATC 681
 QY 165 TyrProAspAlaLeuGluLeuLeuLysThrPheAsnAspProThrValPheAlaAla 184
 DB 682 TATCAAAAGCTTATGAGAGCTCTAAAGAGCTTCAATGATGAGACGTTTATGCTGCA 741
 QY 185 ThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArgLeuThrAspIle 204
 DB 742 ACAGGACATTTGATCTAGAAACAGCAAACTATCTTTTAAACGCGATTCACATATC 801
 QY 205 ArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThrGlyAsnIleLeu 224
 DB 802 CGTTACGATAATGCCCTTTGGGGTGGAGCGTCTGCTCAATCAATTAACAGGTAATTTTA 861
 QY 225 ValCysSerGlyProLeuSerValTyrArgArgGluValValProAsnIleAspArg 244
 DB 862 GTTTGCTCAGGACCATTTAGTATTTATCGACGTGAAGTGATTTATTCCTAACCTAGAGCG 921
 QY 245 TyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAspArgCysLeuThr 264

Db 922 TATATAAAATCAAAATTCCTAGGTTTACCTGTTAGCATTTGGGGATGATCGATGTTTAAACA 981
 QY 265 AsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLysCysIleThrAsp 284
 Db 982 AATTATGCTATTGATTAGGACGACTGCTACCAATCAACAGCTAGATGTGATAGTAT 1041
 QY 285 ValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTyrPheAsnLysSerPhePhe 304
 Db 1042 GTACCTTTCCCAATTAAGAGTTATTTAAAGCAACAAATCGATGGAATAAATCTTTTTT 1101
 QY 305 ArgGluSerIleLeuSerValLysLysIleMetAsnAspProPheValAlaLeuTyrThr 324
 Db 1102 AGAGATCTATTATTTCTTAAATAAATTTCTTAAATCCCATCGTTGCTTATGGACT 1161
 QY 325 IleLeuGluValSerMetPheMetMetLeuValTyrSerValValAspPhePheValGly 344
 Db 1162 ATTTTCGAAGTCTGTTATGTTATGATGTTGTCGAATTTGGAATCTTTTGTGTTAAT 1221
 QY 345 AsnValArgGluPheAspTyrLeuArgValLeuAlaPheLeuValIleIlePheIleVal 364
 Db 1222 CAAGCTATTCAATTAGACCTTATTAACTTTTTCCTTTTATCCATCATCTTTATCGGT 1281
 QY 365 AlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPheLeuLeuSerPro 384
 Db 1282 GCATTATGCTGTAATGTTTATATATATGTCGAACATCTGCTAGTTTGTGTTATCTCT 1341
 QY 385 PheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyrSerLeuPheThr 404
 Db 1342 CTGTATGGAATATTACACTTTGTTGTCCTACAGCCCTAAACCTTTATCTTATCTGACC 1401
 QY 405 IleArgAsnAlaAspTyrPheGlyThrArgLysLysLeu 416
 Db 1402 ATTAAATACGGAATGGGACACGTAAGAGTC 1437

RESULT 2

US-09-146-993-1
 ; Sequence 1, Application US/09146893
 ; Patent No. 6455304
 ; GENERAL INFORMATION:
 ; APPLICANT: Weigel, Paul H.
 ; DeAngelis, Paul L.
 ; Papanstantinou, John
 ; TITLE OF INVENTION: Hyaluronate Synthase Gene and Uses
 ; Thereof
 ; NUMBER OF SEQUENCES: 7
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSES: Arnold, White & Durkee
 ; STREET: P. O. Box 4433
 ; CITY: Houston
 ; STATE: TX
 ; COUNTRY: USA
 ; ZIP: 77210-4433
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/146,893
 ; FILING DATE: 03-Sep-1998
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/270,581
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Parker, David L.
 ; REGISTRATION NUMBER: 32,165
 ; REFERENCE/DOCKET NUMBER: UTSG:161\PAR
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 713-787-1400
 ; TELEFAX: 713-789-2679
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:

; LENGTH: 1512 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 193..1449
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1488..1510
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-146-893-1

Alignment Scores:
 Pred. No.: 4 66e-187 Length: 1512
 Score: 1586.50 Matches: 298
 Percent Similarity: 85.19% Conservative: 53
 Best Local Similarity: 72.33% Mismatches: 60
 Query Match: 73.96% Indels: 1
 DB: 4 Gaps: 1

US-09-469-200D-2 (1-417) x US-09-146-893-1 (1-1512)

QY 5 LysAsnLeuIleThrValValAlaPheSerIlePheTrpValLeuLeuIleTyrValAsn 24
 DB 205 AAAAAAATTAATGTTTATCTTTATTTTGGATATCTATCTGATTATCTAAT 264
 QY 25 ValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeuIleAlaTyr 44
 DB 265 ATGTATCTATTGGAGACA---TCAACTGTGAGGAATTTATGGAGTAATATTAATACCTAT 321
 QY 45 LeuLeuValLysMetSerLeuSerPhePheTyrLysProPheLysGlyArgAlaGlyGln 64
 DB 322 CTAGTTATCAACITGGATTATCTTCTTTATGAGCCATTAAGGAAATCCACATGAC 381
 QY 65 TyrLysValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSerLeuLeuGluThr 84
 DB 382 TATAAAGTGTCTGTGTAATCTCTTATAATGAAGTGGCAGTCATATTAGAAACA 441
 QY 85 LeuLysSerValGlnGlnThrTyrProLeuAlaGluIleTyrValValAspGly 104
 DB 442 CTAAAGTGTGTAGCACAGACCTATCGTATCAACAATTTATATTGTTGATGG 501
 QY 105 SerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgAspThrGlyAspLeu 124
 DB 502 AGTCAACACAGATGCAATACAAATTAATGAAGATGTAATAGAGAGTGGATAT 561
 QY 125 SerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArgHisAlaGlnAla 144
 DB 562 TGTCGAACGTTATCGTTCACCGTTCCTTCTCAATAAAGGAAACGCCATGCTCAAGG 621
 QY 145 TrpAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSerAspThrTyrIle 164
 DB 622 TGGGCATTTGAAGATCTGACGTCAGCTTTTTTAACCGTAGACTCAGACTATATC 681
 QY 165 TyrProAspAlaLeuGluGluLeuLeuLysThrPheAsnAspProThrValPheAlaAla 184
 DB 682 TATCCAAATGCTTAGAAGAACTCTCAAAAAGCTCAATGATGAGACAGTTTATGCTGCA 741
 QY 185 ThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArgLeuThrAspIle 204
 DB 742 ACAGGACATTTGAATGCTAGAACACAGACAACTAATCTATTACGCGACTTACAGATATC 801
 QY 205 ArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThrGlyAsnIleLeu 224
 DB 802 CGTTACGATTAATGCTTTGGGGTGGAGCGTCTCTCAATCAATTAACAGGTATATTTA 861
 QY 225 ValCysSerGlyProLeuSerValTyrArgArgGluValValProAsnIleAspArg 244
 DB 862 GTTTGCTCAGACCACTAGATATTATTCGAGCTGAAGTATTATTCCTACTTAGAGCC 921
 QY 245 TyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspArgCysLeuThr 264

DB 922 TATAAAATCAACATCTCTAGGTTTACCTGTAGCATTTGGGATGATCGATGTTTAAACA 981
 QY 265 AsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLysCysIleThrAsp 284
 DB 982 AATATGCTATTGATTTAGAGCCACTGTCTACCAATCAACAGCTAGATGATCTGAT 1041
 QY 285 ValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTyrAsnLysSerPhePhe 304
 DB 1042 GTACCTTTCCAAATTAAGAGTTATTAAAGCAACAATAATCGATGAATAATCTTTTT 1101
 QY 305 ArgGluSerIleLeuSerValLysIleMetAsnAsnPropheValAlaLeuThrPhe 324
 DB 1102 AGAATCTATATTTCTTTAAATAATTTCTTTTAAATCCCATCGTGCCTATGGACT 1161
 QY 325 IleLeuGluValSerMetMetMetMetValTyrSerValValAspPhePheValGly 344
 DB 1162 ATTTTCGAAGTCGTTATGTTATGATGTTGATGTCGCAATTTGGATCTTTTGTAT 1221
 QY 345 AsnValArgGluPheAspTrpLeuArgValLeuAlaPheLeuValIleIlePheIleVal 364
 DB 1222 CAAGCTATCAATTAGACCTTATTAAACTTTTGGCTTTTATCCATCATCTTTATCGT 1281
 QY 365 AlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPheLeuLeuSerPro 384
 DB 1282 GCTTTATGCTGTAATGCTTATGATGTTATGATGTTGATGTTGATGTTGATGTTG 1341
 QY 385 PheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyrSerLeuPheThr 404
 DB 1342 CTGTATGGAATATATACACTGTTGCTTACAGCCCTTAAACTTATCTTTATGACCC 1401
 QY 405 IleArgAsnAlaAspTrpGlyThrArgLysLysLeu 416
 DB 1402 ATTAAATACGAATGGGAACACGTAATAAAGGTC 1437

RESULT 3

; Sequence 1, Application US/09214808A
 ; Patent No. 6475793
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosenthal, Andre
 ; APPLICANT: Freiberg, Christoph
 ; APPLICANT: Perret, Xavier Philippe
 ; APPLICANT: Broughton, William John
 ; TITLE OF INVENTION: Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic
 ; Patent No. 6475793
 ; TITLE OF INVENTION: Plasmid
 ; FILE REFERENCE: CARP0068
 ; CURRENT APPLICATION NUMBER: US/09/214,808A
 ; CURRENT FILING DATE: 1999-06-22
 ; PRIOR APPLICATION NUMBER: PCT/IB97/00950
 ; PRIOR FILING DATE: 1997-07-10
 ; NUMBER OF SEQ ID NOS: 1
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 536165
 ; TYPE: DNA
 ; ORGANISM: Rhizobium
 US-09-214-808-1

Alignment Scores:
 Pred. No.: 3 47e-43 Length: 536165
 Score: 463.00 Matches: 131
 Percent Similarity: 48.75% Conservative: 64
 Best Local Similarity: 32.75% Mismatches: 185
 Query Match: 21.59% Indels: 20
 DB: 4 Gaps: 10

US-09-469-200D-2 (1-417) x US-09-214-808-1 (1-536165)

QY 27 LeuPheGlyAlaLysGly-----SerLeuSerIleTyrGlyPheLeuLeuIleAlaTyr 44
 DB 157330 CTGCTTGGCAGCGCGCGGTAGCATCTCCTTGTATGAGCAGCTCTCGACGCGCTTAC 157271

QY 45 -----LeuLeuValLysMetSerLeuSerPhePheTyrLysProPhe 58
 Db 157270 AAAGCGATCGAAGCTATATACCGTTTGCACAAACACACAGCGCGGTCAACGCCCGCG 157211

QY 59 LysGlyArgAlaGlyGlnTyrLysValAlaAlaIleProSerTyrAsnGluAspAla 78
 Db 157210 ACCGCTCGGTCGACCGCGAGCGTGTATCGTACCGCTGCTACATGAGATCCG 157151

QY 79 GluSerLeuGluThrLeuLysSerValGlnGlnGlnThrTyrPro----LeuAlaGlu 97
 Db 157150 CGGCGCGCTCGCGCGTCTAGCTTCCATTGCAAGACGAGACTACGCTGAGAGTTCGG 157091

QY 98 IleTyrValValAspAspGlySerAlaAspGluThrGlyLeuLysArgIleGluAspTyr 117
 Db 157090 GTCTAGTGGTGGACGAGCTTCTGCAATCGCAAGCCATCATCTGCTACAGATCAT 157031

QY 118 ValArgAspThrGlyAspLeuSerSerAsnValIleValHisArgSerGluLysAsnGln 137
 Db 157030 TAT-----GGTCCGACCGAGGTTCGCTTTATCTCTG-----ATGCCAAGAAGCTC 156983

QY 138 GlyLysArgHisAlaGlnAlaTrpAlaPheGluArgSerAspAlaAspValPheLeuThr 157
 Db 156982 GGAAGCCGAGGCGGAGATTCGCGATACGGGAATCATCGGAGATTTGGTGTCTAAC 156923

QY 158 ValAspSerAspThrTyrIleTyrProAspAlaLeuGluGluLeuLysThrPheAsn 177
 Db 156922 GTTGACTCGACACGACCATTCGCGCGAGCTAGTCACGAACTTCGCTGAAGATGTAC 156863

QY 178 AspProThrValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeu 197
 Db 156862 AGTCCCGCGTGGCGCGCGTGTGGTTCAGTTCGCGCGCAGCAACCGCAGCACATGG 156803

QY 198 LeuThrArgLeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGln 217
 Db 156802 CTGACGCGGTGATCAGATGAGTCTGGCTCGCTCGCAGGAGAACCGCAGCACAG 156743

QY 218 SerValThrGlyAsnLeuValCysSerGlyProLeuSerValTyrArgArgGluVal 237
 Db 156742 GCTCGCTTTGGAGCGGTATGTGTGTCGCGCGCGTGTGCCATGACCGCGGTCGCA 156683

QY 238 ValValProAsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIle 257
 Db 156682 CTCCATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 156623

QY 258 GlyAspArgCysLeuThrAsnTyrAlaThrAspLeuGly---LysThrValTyrGln 276
 Db 156622 GGGGAAGACCGCCACCTCACATCTCATCTGATGACGCTTTCGAACCGAGTACGTT 156563

QY 277 SerThrAlaLysCysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGln 296
 Db 156562 CGGAAGCCATCGCGCGAGCGGTGTTCCAAACTCGATGGCGCTATCTGCGCCACAA 156503

QY 297 AsnArgTyrAsnLysSerPheArgGluSerIleIleSerValLysLysIleMetAsn 316
 Db 156502 CTGCGCTGGGCGAGCAGCTTTCGGGACACATGCTGCGCTGCGCTGCTGCTGCTGCTGCT 156443

QY 317 ---AsnProPheValAlaLeuThrThrIleLeuGluValSerMetPheMetLeuVal 335
 Db 156442 CTTGATCGCTATCTTACGCTGAGCTGATCGGACAGATCTTGGCTGCTGCTGCTGCTGCT 156383

QY 336 TyrSerValValAspPhePhe-----ValGlyAsnValArgGluPheAspTrpLeuArg 353
 Db 156382 CTCTCGGTCGTGAGCGGCTAGCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 156323

QY 354 ValLeuAlaPheLeuValIleIlePheIleValAlaLeuCysArgAsnIleHisTyrMet 373
 Db 156322 ATCTGT-----ATGATTCGATCATGACAAATGGTCCGCTGC---GGCTGTGCGCGCTT 156272

QY 374 LeuLysHisProLeuSerPheLeuLeuSerProPheTyrGlyValLeuHisLeuPheVal 393
 Db 156271 CGAGCGGAGAGCTGGATCTCTTGGGTTTTCGCTGACACCTCTCTCAACGCTGCTCTC 156212

QY 394 LeuGlnProLeuLysLeuTyrSerLeuPheThrIleArgAsnAlaAspTrpGlyThrArg 413
 Db 156211 CTGTCCCGCTCAAAAGCATATGGCTGTGTGACGTTGAGCAACAGCGACTGTGCTGCGCT 156152

RESULT 4
 ; US-08-675-499A-1
 ; Sequence 1, Application US/08675499A
 ; Patent No. 6492150
 ; GENERAL INFORMATION:
 ; APPLICANT: McDonald, J. A.
 ; APPLICANT: Spicer, A. P.
 ; APPLICANT: Augustin, M. L.
 ; TITLE OF INVENTION: GENE ENCODING HYALURONAN
 ; TITLE OF INVENTION: SYNTHASE
 ; NUMBER OF SEQUENCES: 32
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
 ; STREET: P.O. Box 2938
 ; CITY: Minneapolis
 ; STATE: MN
 ; COUNTRY: USA
 ; ZIP: 55402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/675,499A
 ; FILING DATE: 03-JUL-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Woessner, Warren D
 ; REGISTRATION NUMBER: 30,440
 ; REFERENCE/DOCKET NUMBER: 150.170US1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 612-359-3260
 ; TELEFAX: 612-359-3263
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2947 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; US-08-675-499A-1

Alignment Scores:
 Pred. No.: 5,99e-44 Length: 2947
 Score: 439.00 Matches: 134
 Percent Similarity: 46.65% Conservative: 82
 Best Local Similarity: 28.94% Mismatches: 183
 Query Match: 20.47% Indels: 65
 DB: 4 Gaps: 14

US-09-469-200D-2 (1-417) x US-08-675-499A-1 (1-2947)

QY 7 LeuIleThrValValAlaPheSerIlePheTrpValLeuLeuIleTyrValAsnValTyr 26
 Db 574 CTCTCGGAATCACAGCTGCTATATTGTTGGCTACCATTTTCCAAACAGATAATTAC 633

QY 27 LeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeuLeuIleAlaTyrLeuLeu 46
 Db 634 TACTTC-----TCAATTGGACTGTAGGTCGCTTTTGGCTTCGCTCATCTCATC 681

QY 47 Val-LysMetSerLeuSerPhePhe-----TyrLysProPheLysGlyArg 61
 Db 682 ATCCAAAGCCTCTTTGCTTTTGGTAAACACCGGAAATGAGAGATGCTCTTGAACCCCG 741

```

QY 61 gAlaGlyGlnTyrLys---ValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSe 80
Db 742 ATTAATG-ACAAACGGTAGCACTCGCATCGCTGCCAGCAGAGAGCCCTGACTA 800
QY 80 rLeuLeuGluThrLeuLysSerValGlnGlnGlnThrTyrProLeuAlaGluIleTyrVa 100
Db 801 CTACGGAATGTTTGAATCTGTGAAGAGCTGACCTACCTGGGATTAAAGTCGTGAT 860
QY 100 lValAspAspGlySerAlaAsp-----107
Db 861 GGTCATCGATGGGAACCTCAGACGACGACCTTTACATGATGAGACATATTACCGAAGTAT 920
QY 108 -----GluThrGlyIleLys 112
Db 921 TGGCAGGACAAATCGGCCACGTACATCTGGAAGCAACAACCTTTCATGAAGAGGACCTGG 980
QY 112 sArgIleGluAspThrValArgAspThrGlyAsp-----LeuSerSerAs 127
Db 981 TGACAGAGAAGAGTCCCATTAAGAAAGTTTCAACATGTCAACCAATTGCTGTGTCTAA 1040
QY 127 nValIleValHisArgSerGluLysAsnGlnGlySerHisAlaGlnAlaTrpAlaPh 147
Db 1041 CAAAGATATTGCATCATGTGCAAAATGGGGTGGAAAGAGAGAGTCAATGTACACAGCCT 1100
QY 147 eGluArg-----SerAspAlaAspValPheLeuThrValAspSerAspThrTyrIle 165
Db 1101 CAGAGCACTGGGGCAAGCGTGCATTTATGTACAGGTGTGTGACTCAGATCATGTCTGA 1160
QY 165 rProAspAlaLeuGluGluLeuLeuLysThrPhe---AsnAspProThrValPheAla 184
Db 1161 CCTGTGCTCATCTGGAGATGTGAAGTCTTAGAGAGAGACCTATGTGTGGAGTGT 1220
QY 184 aThrGlyHisLeuAsnValArgAsnGlnThrAsnLeuLeuThrArgLeuThrAspI 204
Db 1221 TGGAGGAGATGCCAGATTTAAACAAGTATGATCTCGGATCTCTTCCCTCAGCAGCT 1280
QY 204 eArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThrGlyAsnIle 224
Db 1281 GAGATACTGGATGGCTTTAATATAGAAGGGCTGCCAGTCTTATTTGGCTGTGTCCA 1340
QY 224 uValCysSerGlyProLeuSerValTyrArgGluValValValProAsnIleAsp 244
Db 1341 GTCATACAGCGTCTCTGGGAATGTACAGAACTCCTGTGTCATGAATTTGTGGAAGA 1400
QY 244 gTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAspArgCysLeu 264
Db 1401 CTGGTACAAATCAGGAATTCATGGTAACCAATGCAGTTTGGTGACGACGAGCACCTTAC 1460
QY 264 rAsnTyrAlaThrAspLeuGly---LysThrValTyrGlnSerThrAlaLysCysIle 283
Db 1461 CAACAGGTTGTGAGTCTGGGCTATGCAACTAAATACAGGCTCGGTCGAAGTCCTTAC 1520
QY 283 rAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrpAsnLysSer 303
Db 1521 TGAATCTCCATAGATATCTGAGATGCTGAACACGACGACCCGATGGAGCAGTCTTA 1580
QY 303 ePheArgGluSerIleSerValLysIleMetAsnAsnProPheValAlaLeuTr 323
Db 1581 CTTCAGAGTGGTGTACAAATGCCATGTGTTTCAACAGCAT-----CACCTGTG 1631
QY 323 pThrIleLeuGluValSerMet-----PheMetMetLeuValTyrSerVa 338
Db 1632 GATGACCTATGAAGCTGTTATCATCGGATCTTTTCCTTTCTTCATTGCCACATCAT 1691
QY 338 lValAspPheValGlyAsnValArgGluPheAspTrpLeuArgValLeuAlaPheLe 358
Db 1692 CCAGCTCTTCTACAGGGGTAAATC-----TGG---AACATCTCTCTCTCTCT 1736
QY 358 uValIleIlePheIleValAlaLeuCysArg---AsnIleHisTyrMetLeuHisPr 377
Db 1737 GTTAACTGTCAGCTAGTGGGTCTCATCACTCATCTTTTGGCAGCTCCCTTAGAGAAA 1796
QY 377 oLeuSerPheLeuLeuSerProPheTyrGlyValLeuHisLeuPheValLeuGlnPro 397

```

```

Db 1797 TATCTCATGCTATTTCATCTCTCTGTTATTCAGTGTATATACATGCTACTTCTCTGC 1856
QY 397 ulysLeuTyrSerLeuPheThrIleAlaGlnAlaAspTrpGlyThr-----ArgLysLy 415
Db 1857 CAAGATGTTTGCATTTGCAATTTGCAACCAATAACAAGCTGGTGGGACATCTGGGAAGAGAC 1916
QY 415 sleuLeu 417
Db 1917 CATTGTT 2923

RESULT 5
US-09-155-768-1
; Sequence 1, Application US/09155768A
; Patent No. 6162908
; GENERAL INFORMATION:
; APPLICANT: SEIKAGAKU CORPORATION
; TITLE OF INVENTION: POLYPEPTIDE OF HYALURONAN SYNTHASE OF
; TITLE OF INVENTION: HUMAN-ORIGIN AND DNA ENCODING THE SAME
; FILE REFERENCE: TOYAM37.001APC
; CURRENT APPLICATION NUMBER: US/09/155,768A
; EARLIER FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: JP 8-084326
; EARLIER FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: JP 8-109663
; EARLIER FILING DATE: 1996-04-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2117
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (149)...(1777)
US-09-155-768-1

```

```

Alignment Scores:
Pred. No.: 2e-35 Length: 2117
Score: 368.50 Matches: 137
Percent Similarity: 42.08% Conserved: 73
Best Local Similarity: 27.45% Mismatches: 182
Query Match: 17.18% Indels: 107
DB: Gaps: 20

US-09-469-200d-2 (1-417) x US-09-155-768-1 (1-2117)

QY 4 LeuLysAsnLeuIleThrValValAlaPheSerIlePheTrpValLeuLeuIleTyrVal 23
Db 98 CTGGCCCGGAGGGTGTGACCATCGCTTCGCCCTGCTC---ATCTGGCCCTCATGACC 154
QY 24 AsnValTyrLeuPheGlyAlaLys-----GlySerLeuSer----- 35
Db 155 TGGGCGCTACCGCCCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 214
QY 36 lIleTyrGlyPheLeuLeuIleAlaTyrLeuLeuValLysMetSerLeuSerPhePheTyr 55
Db 215 CTCTACCGCGGCTCTCTTTCAGCCAC---CTGGTGGGCGAGAGCCTC---TTCGCGTAC 268
QY 56 LysProPheLysGlyArgAlaGlyGlnTyrLys----- 66
Db 269 CTGGAGCACCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 328
QY 67 ---ValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSerLeuLeuGluThrLeu 85
Db 329 AGTGTGGCGTGCACCATCTCGGCTACAGGAGAGACCGCGCTACCTGGCCAGTGCCTG 388
QY 86 LysSerValGlnGlnThrTyrProLeuAla----- 96
Db 389 CGGTCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 448
QY 97 -----GluIleTyrValValAsp----- 102

```

```
Db 449 GGCAACCGCGCGAGGACCTTACATGGTCGACATGTTCCGCGAGGTCCTTCGCTGACGAG 508
QY 103 -----AspGlySer----- 105
Db 509 GACCCGCCACATGAGTGGGACGACAACTACACACAGCCCTGGGAACCCGCGGCGGCG 568
QY 106 ---AlaAspGluThrGlyLeuArgAspThrValArgAspThrGlyAspLeu 124
Db 569 GCGCGCTGGCGCGGAGGCTATCG--GAGGTGAGCGGAGGATCCTGGCGGCGCTG 625
QY 125 SerSerAsnValIleValHisArg-----SerGluLysAsnGlnGly 138
Db 626 GCGATGGAGCGCGTGGTGAGGACTCCGAGTGCCTGTCGTGCGCGCAGCGTGGGCGCG 685
QY 139 LysArgHisAlaGlnAlaTrpAlaPheGluArg-----SerAspAlaAspValPheLeu 156
Db 686 AAGCGCGAGTCATGTACACAGCCTTCAGCGCTCGGAGATTCGGTGGATACGGTGCAG 745
QY 157 ThrValAspSerAspThrTrpIleTyrProAspAlaLeuGluLeuLysThrPhe 176
Db 746 GTCTGTGCTGACGACACAAAGTTGGACCCCATGCGACTGCTGGAGCTCGTGGGTACTG 805
QY 177 Asn---AspProThrValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThr 195
Db 806 GACGAGACCCCGGAGGCTGTGGTGGGATGTGGGATCGGATCCTTAACCTCTGGAC 865
QY 196 AsnLeuLeuThrArgLeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAla 215
Db 866 TCCTGGTCTGAGTCTCAACAGCAGCTGGATCTGGTACCTTCAATGTGGAGCGGT 925
QY 216 AlaGlnSerValThrGlyAsnIleLeuValCysSerGlyProLeuSerValTyrArg 235
Db 926 TGCAGAGTACTTCCACTGTATCTGTCATCGACGCGTCTCTAGGCTATATAGGAAT 985
QY 236 GluValValProAsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProVal 255
Db 986 AACCTCTCCAGCAGTCTTCTGAGCGCTGGTACACAGAGTCTCTGGGTACCCACTGT 1045
QY 256 SerIleGlyAspArgCysLeuThrAsnTyrAlaThrAspLeuGly---LysThrVal 274
Db 1046 ACTTTGGGATGACGGGACCTCACCACCGCATGCTCAGCATGGTATATACCAAG 1105
QY 275 TyrGlnSerThrAlaLysCysIleThrAspValProAspLysMetSerThrTyrLeuLys 294
Db 1106 TACACTCCAGTCCCGCTGCTACTACAGACGCGCTCGTCTCTCTCGGTGCTGAGC 1165
QY 295 GlnGlnAsnArgTyrAsnLysSerPheArgGluSerIle----- 308
Db 1166 CAGCAGACAGCTGTCCAAAGTCTGACTTCCGTGAGTGGCTGTACAAGCGCTCTGGTG 1225
QY 309 -----IleSerValLysIleMetAsnAsp-----ProPheVal 320
Db 1226 CACCGGCAACATGCGTGGATGAGTACAGAGCGGTGCTCTCGGCGCTGTCCCTTCTTC 1285
QY 321 AlaLeuTriPThrIleLeuGluValSerMetPheMetMetLeuValTyrSerValAsp 340
Db 1286 GTGGCGGCGCTGCTGCTGCTG----- 1309
QY 341 PhePheValGlyAsnValArgGluPheAspThrPheArgValLeuAlaPheLeuValIle 360
Db 1310 TTCTACGCGGCGCGCTTGGGCGTGTGTGGGTG-----CTGCTGTC 1354
QY 361 IlePheIleValAlaLeuCysArg---AsnIleHisTyrMetLeuLysHisProLeuSer 379
Db 1355 GTCCAGGGCGTGGCACTGGCCAAAGCGGCGCTTCGCGCGCTGCGGGGTGCTGCGC 1414
QY 380 PheLeuSerProPheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeu 399
Db 1415 ATGTGCTTCTGCTGTACGCGCGCTCTACATGTGTGGCTCTCTGCTGCTCAAGTTC 1474
QY 400 TyrSerLeuPheThrIleArgAsnAlaAspTrpGlyThr-----ArgLysLysLeu 416
Db 1475 CTGGCGCTAGTACCATGACACAGAGTGGCTGGGCGACCTCGGCGCGCGGAGCTG 1531
```

RESULT 6

```
US-08-635-552A-1
; Sequence 1, Application US/08635552A
; Patent No. 6423514
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; TITLE OF INVENTION: Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1769
US-08-635-552A-1
```

```
Alignment Scores:
Pred. No.: 3,55e-35 Length: 2116
Score: 366.50 Matches: 135
Percent Similarity: 42.12% Conservative: 76
Best Local Similarity: 26.95% Mismatches: 179
Query Match: 17.09% Indels: 111
DB: 4 Gaps: 19
```

US-09-469-200D-2 (1-417) x US-08-635-552A-1 (1-2116)

```
QY 4 LeuLysAsnLeuIleThrValValAlaPheSerIlePheTrpValLeuIleTyrVal 23
Db 90 CTGGCCCGGAGGGTGTGCTGACCATCGCTTCGCCGTG-----CTCATCTGGGCTC 140
QY 24 AsnValTyrLeuPheClyAlaLysGlySerLeuSer----- 35
Db 141 ATGACCTGGCGCTACGCGCGGCTGGCTGGCTCGCATCGCTAGCGCTCTCTGGCC 200
QY 36 -----IleTyrClyPheLeuLeuIleAlaTyrLeuLeuValLysMetSerLeuSerPhe 53
Db 201 TTCGGCTCTACGGGCTTCTCTTCAGCGGCAC---CTGGTGGCGCAGAGCTC---TTC 254
QY 54 PheTyrLysProPheLysGlyArgAlaGlyGlnTyrLys----- 66
Db 255 CGGTACCTGGAGCAGCGGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 314
QY 67 -----ValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSerLeuLeuGlu 83
Db 67 -----ValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSerLeuLeuGlu 83
```


QY	378	LeuSerPheLeuLeuSerProPheTyrGlyValLeuHisLeuPheValLeuGlnProLeu	397
Db	1401	CTGGCCATGGGTTCTTCGCTCTACGGCGCCCTCATATGTGTGGCTCTCGCTGCC	1460
QY	398	LysLeuTyrSerLeuPheThrIleArgAsnAlaAspIrrpGlyThr-----ArgLysLys	415
Db	1461	AGTTTCCTGGCGGTAGTACCATGTACACAGATGGCTGGGGCACCTCGGGCCGGCGGAAG	1520
QY	416	Leu 416	
Db	1521	CTG 1523	
RESULT 7			
US-08-867-030B-5			
: Sequence 5, Application US/08867030B			
: Patent No. 5948900			
: GENERAL INFORMATION:			
: APPLICANT: Yother et al.			
: TITLE OF INVENTION: Streptococcus pneumoniae			
: TITLE OF INVENTION: Capsular Polysaccharide Genes and Flanking Regions			
: NUMBER OF SEQUENCES: 20			
: CORRESPONDENCE ADDRESS:			
: ADDRESSEE: Benjamin Aaron Adler, Ph.D., J.D.			
: STREET: 8011 Candle Lane			
: CITY: Houston			
: STATE: TX			
: COUNTRY: USA			
: ZIP: 77071			
: COMPUTER READABLE FORM:			
: MEDIUM TYPE: 3.5 inch floppy			
: COMPUTER: Apple Macintosh			
: OPERATING SYSTEM: Macintosh			
: SOFTWARE: Microsoft Word for Macintosh			
: CURRENT APPLICATION DATA:			
: APPLICATION NUMBER: US/08/867,030B			
: FILING DATE: June 2, 1997			
: CLASSIFICATION: 435			
: PRIOR APPLICATION DATA:			
: APPLICATION NUMBER: 08/243,546			
: FILING DATE: May 16, 1994			
: ATTORNEY/AGENT INFORMATION:			
: NAME: Benjamin Aaron Adler, Ph.D., J.D.			
: REGISTRATION NUMBER: 35,423			
: REFERENCE/DOCKET NUMBER: D5923			
: TELECOMMUNICATION INFORMATION:			
: TELEPHONE: (713) 777-2321			
: TELEFAX: (713) 777-6908			
: INFORMATION FOR SEQ ID NO: 5:			
: SEQUENCE CHARACTERISTICS:			
: LENGTH: 4951 bp			
: TYPE: nucleic acid			
: STRANDEDNESS: single-stranded			
: TOPOLOGY: linear			
: MOLECULE TYPE:			
: DESCRIPTION: other nucleic acid			
: HYPOTHETICAL: no			
: ANTI-SENSE: no			
: FRAGMENT TYPE:			
: ORIGINAL SOURCE:			
: IMMEDIATE SOURCE:			
: POSITION IN GENOME:			
US-08-867-030B-5			
Alignment Scores:			
Pred. No.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
DB:			
US-09-469-200B-2 (1-417) x US-08-867-030B-5 (1-4951)			
Length: 4951			
Matches: 114			
Conservative: 78			
Mismatches: 103			
Indels: 65			
Gaps: 19			

```

Qy 349 e-----AspTrpLeuArgValLeuAlaPheLeuValIleIlePh 362
Db 2306 AACTACAATTGTTATATACAGCTTCATGGTGGGAATAATTATTTATATGTTCTTTGGGAAT 2365
Qy 362 eIleValAlaLeu---CysArgAsnIleHisIyrMet-----LeuLysHisProLeuSe 379
Db 2366 GATTTTACGTTTGGAGGAGAACTTTAAAGCTAATGCTAGTAATGAGTGGTATTATATCT 2425
Qy 379 rPheLeuLeuSerProPheTyrGlyValLeuHisLeuPheValLeuInProLeuLysLe 399
Db 2426 ATTCTTATATCTGTTTATTAATCGTTTGTAGTATA---ATTATGTCCTCATTAGGCT 2482
Qy 399 uTyrSerLeuPheThrIleArgAsnAlaAsp-----TrpGlyThrArg 413
Db 2483 ATTAGCACTT-----ATGAGATGTTCTGATGATTAGGGTGGGAAC TAGG 2528

RESULT 8
PCT-US95-06119-5
; Sequence 5, Application PC/TUS9506119
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE CAPSULAR
; TITLE OF INVENTION: POLYSACCHARIDE GENES AND FLANKING REGIONS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06119
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: US 08/243,546
; FILING DATE: 16-MAY-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: AMCY018P--
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4951 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
PCT-US95-06119-5

Alignment Scores:
Pred. No.: 3,76e-19 Length: 4951
Score: 242.50 Matches: 114
Percent Similarity: 43.84% Conservative: 78
Best Local Similarity: 26.03% Mismatches: 183
Query Match: 11.31% Indels: 65
DB: 5 Gaps: 19

US-09-469-200D-2 (1-417) x PCT-US95-06119-5 (1-4951)

```


QY 338 ValValAspPheVal ----- 343
 Db 997 ATA-----TTCTTTTGATTTCCGTTAGTCTTTTTTGCTGGATTATTCGGCATTAATGTA 1050
 QY 344 -----GlyAsnValGluPheAspTrpLeuArgValLeuAlaPheLeuValIleIle 361
 Db 1051 TATCAAGGAACACTACGAACAGATTATATTACTTTTACATTCAGTTTGTGTGTCATCTGC 1110
 QY 362 Phe-----IleValAlaLeuCys ----- 367
 Db 1111 TTTGAATGGTCGCGAGGATTTTCCAGTCTCGTAGCTCCGCTGATCTAGATGATCGAGCA 1170
 QY 368 ArgAsnIleHisTyrMetLeuLysHisProLeuSerPheLeuLeuSerProPheTyrGly 387
 Db 1171 AGAAATTTGAAATACCTTCTTTTCGCCACCACTATATATGCTTTTG -----TTCTGG 1221
 QY 388 ValLeuHisLeuPheValLeuGlnProLeuLysLeuTyrSerLeuPheThrIleArgAsn 407
 Db 1222 ATAGTGACGGCATACGATGTCGACGACTTCATTCAGCAGCTCAAAACGATTTTAGT 1281
 QY 408 AlaAspTrpGlyThrArgLys 414
 Db 1282 TATGGAACGGGCATGTGAA 1302
 RESULT 10
 US-09-328-352-2963
 ; Sequence 2963, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINET
 ; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 2963
 ; LENGTH: 1260
 ; TYPE: DNA
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-2963
 Alignment Scores:
 Pred. No.: 1,06e-19 Length: 1260
 Score: 239,00 Matches: 101
 Percent Similarity: 41,49% Conservative: 72
 Best Local Similarity: 24,22% Mismatches: 168
 Query Match: 11,14% Indels: 76
 DB: Gaps: 17
 US-09-469-200D-2 (1-417) x US-09-328-352-2963 (1-1260)
 QY 40 LeuLeuIleAlaTyrLeuLeuValLysMetSerLeuSerPhePheTyrLys----- 56
 Db 61 TTGTGTAGTCTGGACATGGATGGTAGAGCTCTCTGGTTTCTTTAAAGGGAATAT 120
 QY 57 -----ProPheLysGlyArgAlaGlyGlnTyrLysValAlaIleIle 71
 Db 121 CATGACGACCAACTGCTGCTGAGCCATCAAGTGAAGG-----TGCAGTATTATTAT 171
 QY 72 ProSerTyrAsnGluAspAlaGluSerLeuLeuGlnThrLeuLysSerValGlnGlnGln 91
 Db 172 CCTCTGTTTAAATGAAGAGCTCAG---GTCCGACAAACGATTCTGTTATGCTTACAGACA 228
 QY 92 ThrTyrProLeuAlaGlnIleTyrValValAspAspGlySerAlaAspGluThrGlyIle 111
 Db 229 AAATATCCGAATTTGAAGTCTTGTGTTAAATGACGGAAGTAGTCAGCAGCTGCA--- 285
 QY 112 LysArgIleGluAspTyrValArgAspThrGlyAspLeuSerSerAsnValIleValHis 131
 Db 286 GAAATCTTGAATCGACGCTCAAGACGCCGCGATTAGA-----GTTGGCAC 336
 QY 132 ArgSerGluLysAsnGlnGlyLysArgHisAlaGlnAlaTrpAlaPheGluArgSerAsp 151

```
Db 337 CTGCGAGAA---AACCAAGTAAGGCGAGTCTTAAGATCGGAGTGTAGTAAGTAAA 393
QY 152 AlaSpValPheLeuThrValAspSerAspThrTyrIleTyrProAspAlaLeuGluGlu 171
Db 394 TAGTAATCTAGTGTATGACGAGACGCTTTATACATCCTCATGCAGTACTTGG 453
QY 172 LeuLeuLysThrPhe---AsnAspProThrValPheAlaAlaThrGlyHisLeuAsnVal 190
Db 454 CTTATCAGCCCTTTTAAATTTCCAGAAATGGTCTTAACCTGGAATCCAGAAAT 513
QY 191 ArgAsnArgGlnThrAsnLeuLeuThrArgLeuThrAspIleArgTyrAspAsnAlaPhe 210
Db 514 TTAATTCGCG---TCAAGTATTTAGTAAATTAACAAGTTGGCGAGTCTCTTCATATT 570
QY 211 GlyValGluArgAlaAlaGlnSerValThrGlyAsnIleLeuValCysSerGlyProLeu 230
Db 571 GGTCTGATTAAGCCGCGCAGACATATGCGCGTATTTTACAGTATCAGGTGAAT 630
QY 231 SerValTyrArgArgGluValValProAsnIleAspArgTyrIleAsnGlnThrPhe 250
Db 631 GCAGCATTTAGAAAGCCGCTTTGGTT----- 657
QY 251 LeuGlyIleProValSerIleGlyAspArgCysLeuThr----- 264
Db 658 -----CGCGTGGCGCTTTGGTCTGTATGACAAATATACGGAAGATATTGATCTCT 708
QY 265 ---AsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLysCysIleThr 283
Db 709 TGGAACTCCAAATGGACCATGGACATTCATATATTCCTCAAGCACCTTTGTTATATC 768
QY 284 AspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrpAsnLysSerPhe 303
Db 769 TATATGCGGAAACCTTTAAAGGCCCTTTGGAACAGCGTTTACGCTGGCACAAGTGG 828
QY 304 PheArgGluSerIleIleSerValLysLysIleMetAsnAsnPropheValAlaLeuIrp 323
Db 829 GTAGAGGTATTATTAGATATATCCGAAATGTTCAAACTCCGCTTGCCTGATGTGG 888
QY 324 ThrIleLeu-----GluValSerMetPheMetLeu 334
Db 889 CCGGTAATGCTCGAAGCTTAATCAGTATTATTGGTCATATGTCATGATGATGATTTT 948
QY 335 ValTyrSerValValAspPheValGlyAsnValArgGluPhe----- 349
Db 949 ATTCTATTTTTGTGGGCAATTTGTTGCAATTTACCGCAACAAATTCAAATTAATTC 1008
QY 350 -----AspTrpLeuArgValLeu-----AlaPheLeuValIle 360
Db 1009 ATGCGCGAGTGTATGCGGTGATATAGGTGGAAATGTCTTGTTCATTTCTAGTAAGT 1068
QY 361 IlePheIle-----ValAlaLeuCysArgAsnIleHisTyrMetLeu 374
Db 1069 TTATGGATAGACCATGTTATGATCGTGGTGGCTTTTGAAGAAATTTTGGTGAAT 1128
QY 375 LysHisProLeuSerPhe---LeuLeuSerPropheTyrGlyValLeuHisLeuPheVal 393
Db 1129 TGTATCATCATTTTTCGTTACTGACGTATTATACCACTGTT-----GTT 1176
QY 394 LeuGlnProLeuLysLeuTyrSerLeuPheThrIleArgAsnAlaAspTrp 410
Db 1177 GCAGTACCCAACTATATTATAC-----ACCAAAACAGTCTCGTGG 1221
```

RESULT 11

```
US-09-134-001C-1214
; Sequence 1214, Application US/09134.001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134.001C
```

```
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1214
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1214

Alignment Scores:
Pred. No.: 3,82e-19 Length: 1257
Score: 234.50 Matches: 89
Percent Similarity: 47.09% Conservative: 97
Best Local Similarity: 22.53% Mismatches: 152
Query Match: 10.93% Indels: 57
DB: 4 Gaps: 16

US-09-469-200D-2 (1-417) x US-09-134-001C-1214 (1-1257)

QY 34 LeuSerIleTyrGlyPheLeuLeuIleAlaTyrLeuLeuValLysMet----- 49
Db 19 ATGCATCTGATTTAACTTTTCTATCTCCATTTTATGTCATTTTACTGGATAGTA 78
QY 50 ---SerLeuSerPhePheTyr-----LysProPhe-----Lys 59
Db 79 GGATCGATTTACTATTATTTTATTAAGAAAAACCCCTTTAATGATCATCTTGTAGTAAA 138
QY 60 GlyArgAlaGlyGlnTyrLys---ValAlaAlaIleIleProSerTyrAsnGluAspAla 78
Db 139 TCTGAACATCAACAGTTGAGGCACTCTCTTTTATTTAGTCTGCTACATGAA---AGT 195
QY 79 GluSerLeuLeuGluThrLeuLysSerValGlnGlnThrTyrProLeuAlaGluIle 98
Db 196 GAAACAGTTCAAGACACGCTTCTCTAGTGTATTTATCTCTAGATATATCTGTAAGAAAT 255
QY 99 TyrValValAspAspGlySerAlaAspGluThrGlyIleLysArgIleGluAspTyrVal 118
Db 256 ATCATTTATCAATGATGGAAGTTCGTGATAATCTGCT---GAAATCATCTATGACTTCAAG 312
QY 119 ArgAspThrGly-----AspLeuSerSerAsnValIleValHisArgSerGlu 134
Db 313 AAAATCATCATGATTTAAATTTTGTGACCTC-----GAA 345
QY 135 LysAsnGlnGlyArgHisAlaGlnAlaTrpAlaPheGluArgSerAspAlaAspVal 154
Db 346 GTCATAGAGGTAAAGCTTAATGCACTCAATGAGGAATCAACAGCATCTTACGAATAT 405
QY 155 PheLeuThrValAspSerAspThrTyrIleTyrProAspAlaLeuGluLeuLeuLys 174
Db 406 GTTATGTGTTTACATGCTGACATCTGATGATGACGATCGCTTTTATATGATGAA 465
QY 175 ThrPhe---AsnAspProThrValPheAlaAlaThrGlyHisLeuAsnValArgAsnArg 193
Db 466 GACTTTTAAAAAGATCCAAAAATTAGCGCAGCTACAGGTAATCCAGTATTCGTATATAA 525
QY 194 GlnThrAsnLeuLeuThrArgLeuThrAspIleArgTyrAspAsnAlaPheGlyValGlu 213
Db 526 ---AGTTCATTTTAGGAAAAATACAGCATTAATGAATATGCAAGTATTATTGGTGTATC 582
QY 214 ArgAlaAlaGlnSerValThrGlyAsnIleLeuValCysSerGlyProLeuSerValTyr 233
Db 583 AAGCAAGTCAATCTCTTGCAGGCAATCAATACTATTTCAGGTGTTTTCACACTATT 642
QY 234 ArgArgGluValValProAsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIle 253
Db 643 AAAAAAGTGCACTC-----AAGATGATAGTATTGGGATACCTGACATGATTAAG 696
QY 254 ProValSerIleGlyAspArgCysLeuThrAsnTyrAlaThrAspLeuGlyLysThr 273
Db 697 GATATTGCTGTTTCAATGAAACCTCATCTCTTTTGTATTAC-----GAAAT 741
```

QY 274 ValTyrGlnSerThrAlaLysCysIleThrAspValProAspLysMetSerThrTyrLeu 293
 Db 742 AAGTACGAACACCGTCTCTATCGTCTGATGTTAGTGGCTCAACACTATAGTGGTTATGG 801
 QY 294 LysGlnGlnAsnArgTyrPhe-----AsnLysSerPhePheArgGluSerIleIle 309
 Db 802 AACAAAGGGTTCGATGGGCTCAAGCGGGCATGAAGTACTTTTAAAGAGACTTTTGGCCA 861
 QY 310 SerValLysLysIleMetAsnAsnProPheValAlaLeuTyrThrIleLeuGluValSer 329
 Db 862 ACAATTAACAACTAAGAAATTAATCACTATATATTTAATGTTTGAACAAATCGCATGATT 921
 QY 330 MetPheMetMetLeuValTyrSerValValAspPheVal-----GlyAsnValArg 347
 Db 922 ACATGGGTCTACATCGTACTGATGTTATTTATCTTTTATCTTTTATAGTAATCACAGCAACATCTTA 981
 QY 348 GluPheAspTyrLeuArg-----ValLeuAla 356
 Db 982 GATTACACATATTAATAATATAGTTTTCAMCTTTTTCATCTTTTACGATGACC 1041
 QY 357 PheLeuValIleIle-----PheIleValAlaLeuCysArgAsnIleHisTyrMetLeuLys 375
 Db 1042 TTTATCAATATATCAATATACAGTTTTCCTTATTTATGACAGTCTGCTACGAAAGAAA 1101
 QY 376 HisProLeuSerPheLeuLeuSerProPheTyrGlyValLeuHis 390
 Db 1102 AATATAGTTGGCTGATATATTTTAAAGTTGGTATCCAAACGTTATAC 1146

RESULT 12

US-09-328-352-1781
 ; Sequence 1781, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 1781
 ; LENGTH: 1263
 ; TYPE: DNA
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-1781

Alignment Scores:

Pred. No.: 1.86e-18 Length: 1263
 Score: 229.00 Matches: 90
 Percent Similarity: 44.13% Conservative: 79
 Best Local Similarity: 23.50% Mismatches: 156
 Query Match: 10.68% Indels: 58
 Db: 4 Gaps: 14

US-09-469-200d-2 (1-417) x US-09-328-352-1781 (1-1263)

QY 24 AsnValTyrLeuPheGlyAlaLysGlySerLeuSerIle-----TyrGlyPhe 39
 Db 1 ACATGCAATCTTTGCTGCTATATGTCGCATGCTTTTAAATTCGTTTCTATTATCCA 60
 QY 40 LeuLeuIleAlaTyrLeuValLysMetSerLeuSerPhePheTyrLysProPheLys 59
 Db 61 TTATTTATGCTGATTTATGATGATGTCGCGCATTTTCTATTGGAAG-----GAA 114
 QY 60 GlyArgAlaGlyGlnTyr-----LysValAlaAla 69
 Db 115 CGTAAGCTCCGCTATGATGATGACCGAGCTCCATAGAAAGCTATCTTAAGTTCGCTGTG 174
 QY 70 HierleProSerTyrAsnGluAspAlaGluSerLeuLeuGluThrLeuLysSerValGln 89
 Db 175 CITATTACTTGTTTAAATGAA---GGGGATAATGACAGGAACGATTACTATGCTTTA 231

QY 90 GlnGlnThrTyrProLeuAlaGluIleTyrValValAspAspGlySerAlaAspGluThr 109
 Db 232 AAATTTGGATTATCTCTCACTTTTGAAGTCAATTTGCGATTAATGATGCGAGTTCGGATAATACA 291
 QY 110 GlyIleLysArgIleGluAspTyrValArgAspThrGlyAspLeuSerSerAsnValIle 129
 Db 292 GGT-----CAAGTGTGGTTCGTTAGCTAGTGAACATGAAAACTCGT-----GTT 339
 QY 130 ValHisArgSerGluLysAsnGlnGlyLysArgHisAlaGlnAlaTyrAlaPheGluArg 149
 Db 340 GTTCACTTTCACAA---AACCAAGGTAAAGCCATCGGGTTGCAAGCTGGAAGCTTAATG 396
 QY 150 SerAspAlaAspValPheLeuThrValAspSerAspThrTyrIleTyrProAspAlaLeu 169
 Db 397 ACCGATCCGAGGTTTTTAATGGTATTGATGGTGAATCTTACTTGACCCGATCGACGA 456
 QY 170 GluLeuLeuLeuLysThrPhe---AsnAspProThrValPheAlaAlaThrGlyHisLeu 188
 Db 457 AAGTGGATGGTTCGTCATTTTCTCAAAATCGAGCTGTTCGCCCTGACGGGTAAACCCG 516
 QY 189 AsnValArgAsnArgGlnThrAsnLeuLeuThrArgLeuThrAspIleArgTyrAspAsn 208
 Db 517 CGTATCGTACCGCTCACTCACT---TTATAGTGTCTATTCAGTTCGTTGAGTTCCTCG 573
 QY 209 AlaPheGlyValGluArgAlaAlaGlnSerValThrGlyAsnIleLeuValCysSerGly 228
 Db 574 ATTCTCGCATGATTAAACGTGCTCAGCTACCTTTGGCGCTTTATTTACCGTATCTGTG 633
 QY 229 ProLeuSerValTyrArgGluValValProAsnIleAspArgTyrIleAsnGln 248
 Db 634 GTTATTACGGCTTTTCGTAAGACGGCTGT-----CACCA 669
 QY 249 ThrPheLeuGlyLeuProValSerIleGlyAspArgCysLeuThrAsnTyrAlaThr 268
 Db 670 GTGCACTATGTCACCTAATATATTTGACCGAGACATCGATATTACTTGGAAATTCGA 729
 QY 269 AspLeuGlyLysThrVal---TyrGlnSerThrAlaLysCysIleThrAspValProAsp 287
 Db 730 CGTCCGGCTGGACATTCGTTTGAACCGAATCGCTTGGTTGGATTTTAAATGCTGAA 789
 QY 288 LysMetSerThrTyrLeuLysGlnGlnAsnArgTyrAsnLysSerPheArgGluSer 307
 Db 790 ACTTTAAACGGTTTGTGAAACAAACGTTTACGTTGGCAATGGGTGGAGCAGGATTA 849
 QY 308 IleIleSerValLysLysIleMetAsnAsnProPheValAlaLeuTyrThrIleLeuGlu 327
 Db 850 ATTAAAAATATGATGTTGCTCACTAAACCAAAATTTGAAATTTTATGGCTTTAATG--- 906
 QY 328 ValSerMetMetLeuValTyrSerValValAspPheValGlyAsnValArg 347
 Db 906 ----- 906
 QY 348 GluPheAspTyrLeuArgValLeuAlaPheLeuValIleIlePheIleValAlaLeuCys 367
 Db 907 ---TTGAGCTTTGCTTAACCTTTAGCTGTCATATTAAATGCTGGGAGTTCGTTGTTG 963
 QY 368 ArgAsnIleHisTyrMetLeuLysHisPro---LeuSerPheLeuLeuSerProPhe--- 385
 Db 964 TGGTGGTACATTTTATTTTACGGTCCCTGCTCATAGCCGTTGTGAGCTCTCATTTTA 1023
 QY 386 ---TyrGly 387
 Db 1024 CGTATGGT 1032

RESULT 13

US-08-961-527-86/c
 ; Sequence 86, Application US/08961527
 ; Patent No. 6420135
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 391
 ; CORRESPONDENCE ADDRESS:

```

RESULT 14
US-09-499-203-16
; Sequence 16, Application US/09499203
; Patent No. 6570065
; GENERAL INFORMATION:
; APPLICANT: KOSSMANN, Jens
; APPLICANT: WELSH, Thomas
; APPLICANT: QUANZ, Martin
; APPLICANT: KNUTH, Karola
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Alternansucrase
; FILE REFERENCE: 147-196P
; CURRENT APPLICATION NUMBER: US/09/499,203
; CURRENT FILING DATE: 2000-02-08
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 16
; LENGTH: 4066
; TYPE: DNA
; ORGANISM: Leuconostoc mesenteroides
US-09-499-203-16

Alignment Scores:
Pred. No.:      8,74e-10      Length:      4066
Score:          166.00      Matches:      78
Percent Similarity: 41.64%      Conservative: 64
Best Local Similarity: 22.87%      Mismatches: 122
Query Match:      7.74%      Indels:      77
DB:              4          Gaps:      15

US-09-469-200D-2 (1-417) x US-09-499-203-16 (1-4066)

QY 21 lIeTyValasnValTyLeu-----PheGlyAlaLysGly-SerLeuSerIle-- 36
|||||:|||||: :|||:|||||: :|||:|||||: :|||:|||||: :|||:|||||:
Db 3117 ATTATACATCAATTACTTACGTAGGTGTTTGGTTCAGTTGGTCTATTCGCGTTG 3176

QY 37 -----TyGlyPheLeuLeuIleAlaTyLeuLeuVallySerSe 50
|||||:|||||: :|||:|||||: :|||:|||||: :|||:|||||: :|||:|||||:
Db 3177 TTGTGGAAATTTCCACGCTATTATTCAATCTAGTT-----TTGATAAAACATC 3227

QY 50 rLeu-----SerPheTyLeuTyProPheLysGlyArgAlaGlyGlnTyLysValAl 68
|||||:|||||: :|||:|||||: :|||:|||||: :|||:|||||: :|||:|||||:
Db 3228 AAAAAAGCATCAAGTGATGCCAAGAACCAITGAAA-----TTACCAATGCTATC 3275

QY 68 aAlaIleIleProSerTyraSngLuaspAlaGluSerLeuLeuGluThrLeuLysSerVa 88
|||||:|||||: :|||:|||||: :|||:|||||: :|||:|||||: :|||:|||||:
Db 3276 TATGTGATTAACCTATATATAAT--TCAGGCAACAACATATACCAAGTGATTCAGTCGAT 3332

QY 88 lGlnGlnGlnThrTyPro-----LeuAlaGluIleTyValValaspGlySerAl 106
|||||:|||||: :|||:|||||: :|||:|||||: :|||:|||||: :|||:|||||:
Db 3333 TAGTCGAATCAACCTATCCAAAACAATTATCCAAATTAITGCCGTGATTAATCAAGCAC 3392

QY 106 aspGluThrGlyIleLysArgIleGluAspTyValArgAspThrGlyAspLeuSerSe 126
|||||:|||||: :|||:|||||: :|||:|||||: :|||:|||||: :|||:|||||:
Db 3393 TGAT-----AAAG 3401

QY 126 rAsnValIleValHisArgSerGlu----- 134
|||||:|||||: :|||:|||||: :|||:|||||: :|||:|||||: :|||:|||||:
Db 3402 TTTTACAGCTGTTTATCAGGCTCAAGCTGATTTCCCAATCAATAGATCGATGGATGAA 3461

QY 135 -lysAsnGlnGlyLysArgHisAlaGlnAlaTrpAlaPheGluArgSerAspAlaSpVa 154
|||||:|||||: :|||:|||||: :|||:|||||: :|||:|||||: :|||:|||||:
Db 3462 CACAGATCAAGGTAAAGCGCGTCACATAAATGCGCTATTATATAGATATGGGCAATA 3521

QY 154 lPheLeuThrValaspSerAspThrTyIleTyProAspAlaLeuGluGluLeuLeuLy 174
|||||:|||||: :|||:|||||: :|||:|||||: :|||:|||||: :|||:|||||:
Db 3522 TATTTATTAATTTGGATACTGATGGTTGGTTAGACCAACCAATGCTTTAAAGCGTTTGCT 3581

QY 174 sThrPhe-----AsnAspProThrValPheAlaAlaThrGlyHisLeu----- 188
|||||:|||||: :|||:|||||: :|||:|||||: :|||:|||||: :|||:|||||:
Db 3582 TTATTTTGAAATCACTCAGAAATTCATGTAGCAACTGGTAGCATCTGCACACAAAAA 3641

QY 189 -----AsnValArgAsnArgGlnThrAsnLeuLeuThrArgLeuThrAspIleAr 205
|||||:|||||: :|||:|||||: :|||:|||||: :|||:|||||: :|||:|||||:
Db 3642 ATGATTTCAAAAACGCCAAATATATGGCTTAATTTGCTACAGTTAAACGAATATTTGA 3701

QY 205 gTyraSpAsnAlaPheGlyValGluArgAlaAlaGlnSerValThrGlyAsnIleLeuVa 225
|||||:|||||: :|||:|||||: :|||:|||||: :|||:|||||: :|||:|||||:
Db 3702 GTATGCACATCATCTTATATCAGCGCGTAGTATTGAAACCGTGGTAACTGTTGTTCC 3761

QY 225 lCysSerGlyProLeuSerValTyArgArgGluValValProAsnIleAspArgTy 245
|||||:|||||: :|||:|||||: :|||:|||||: :|||:|||||: :|||:|||||:
Db 3762 AATGTCAGGTGCATTTTCAGCGTTTACGCTGATGATTAGTT----- 3804

QY 245 rIleAsnGlnThrPheLeuGlyIleProValSerIleCysAspArgCysLeuThr-- 264
|||||:|||||: :|||:|||||: :|||:|||||: :|||:|||||: :|||:|||||:
Db 3805 -----CAGACATTATGTATATGTTGCACACCGTTGGTGAAGATACCTGATGACAT 3857

QY 265 -AsnTyraIleThrAspLeuGlyLysThrVal---TyGlnSerThrIleLysCysIleTh 283

```


Search completed: October 2, 2003, 16:14:52
Job time : 269 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2003, 12:56:43 ; Search time 334 Seconds
(without alignments)
3184,760 Million cell updates/sec

Title: US-09-469-200D-2

Perfect score: 2145

Sequence: 1 MRTLKLNITVAFSIFWLL.....KLYSLFTIRNADGTRKLL 417

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1708419 seqs, 1275431651 residues

Total number of hits satisfying chosen parameters: 3416838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framet_p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09469200/runat_02102003_090349_22849/app_query.fasta_1.583
-DB=Published Applications_NA -QPM=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blotum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR_SCORE=ptc -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09469200.@CGN.1.1.221@runat_02102003_090349_22849
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:**

1: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2.6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2.6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2.6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2.6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2.6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2.6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2.6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2.6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2.6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
11: /cgn2.6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2.6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2.6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2.6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2.6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
16: /cgn2.6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
17: /cgn2.6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2145	100.0	1254	10	US-09-879-959-1 Sequence 1, Appli

2	2145	100.0	1254	14	US-10-011-768B-1	Sequence 1, Appli
3	2145	100.0	1254	14	US-10-011-771B-1	Sequence 1, Appli
4	2145	100.0	1254	14	US-10-172-527-1	Sequence 1, Appli
5	2138	99.7	1251	12	US-10-326-185-1	Sequence 1, Appli
6	1976	92.1	5158	12	US-10-326-185-108	Sequence 108, App
7	1695.5	79.0	1251	12	US-10-326-185-102	Sequence 102, App
8	1695.5	79.0	3466	14	US-10-172-527-11	Sequence 11, Appl
9	1586.5	74.0	1257	12	US-10-326-185-92	Sequence 92, Appl
10	1586.5	74.0	1440	14	US-10-172-527-13	Sequence 13, Appl
11	1586.5	74.0	1512	14	US-10-124-222-1	Sequence 1, Appli
12	1586.5	74.0	1512	14	US-10-117-795-1	Sequence 1, Appli
13	977.5	45.6	1200	14	US-10-172-527-17	Sequence 17, Appl
14	463	21.6	536165	11	US-09-939-964-1	Sequence 1, Appli
15	440	20.5	1665	11	US-09-902-939-3	Sequence 3, Appli
16	439	20.5	1859	11	US-09-902-939-2	Sequence 2, Appli
17	439	20.5	4194	14	US-10-262-526-3	Sequence 3, Appli
18	433.5	20.2	3003	11	US-09-918-624B-64	Sequence 64, Appl
19	433.5	20.2	3003	14	US-10-262-526-1	Sequence 1, Appli
20	433.5	20.2	4018	12	US-09-814-353-19942	Sequence 19942, A
21	420.5	19.6	3387	14	US-10-084-817-184	Sequence 184, App
22	369	17.2	1317	14	US-10-156-761-4447	Sequence 4447, Ap
23	369	17.2	9025608	14	US-10-156-761-1	Sequence 1, Appli
24	367	17.1	1752	11	US-09-902-939-1	Sequence 1, Appli
25	366.5	17.1	2116	13	US-10-043-523-1	Sequence 1, Appli
26	343.5	16.0	1740	14	US-10-011-768B-7	Sequence 7, Appli
27	343.5	16.0	1740	14	US-10-011-771B-7	Sequence 7, Appli
28	336.5	15.7	1740	10	US-09-879-959-8	Sequence 8, Appli
29	336.5	15.7	1740	14	US-10-172-527-8	Sequence 8, Appli
30	245	11.4	1380	14	US-10-172-527-15	Sequence 15, Appl
31	242.5	11.3	7430	8	US-08-781-986A-260	Sequence 260, App
32	241.5	11.3	2370	14	US-10-156-761-4813	Sequence 4813, Ap
33	234.5	10.9	4500	10	US-09-771-003-1	Sequence 1, Appli
34	233.5	10.9	570	10	US-09-880-107-3449	Sequence 3449, Ap
35	213	9.9	2250	14	US-10-156-761-5099	Sequence 5099, Ap
36	195.5	9.1	13884	10	US-09-070-927A-341	Sequence 341, App
37	187.5	8.7	17276	10	US-09-870-759-83	Sequence 83, Appl
38	187.5	8.7	17276	12	US-09-751-708A-83	Sequence 83, Appl
39	184.5	8.6	3498	10	US-09-801-368-75	Sequence 75, Appl
40	175	8.2	987	12	US-09-769-744A-47	Sequence 47, Appl
41	173	8.1	1152	10	US-09-974-300-1802	Sequence 1802, Ap
42	171.5	8.0	3396	10	US-09-801-368-71	Sequence 71, Appl
43	169	7.9	2610	12	US-10-246-330-1	Sequence 1, Appli
44	164.5	7.7	490	11	US-09-918-995-27210	Sequence 27210, A
45	160	7.5	2892	10	US-09-801-368-73	Sequence 73, Appl

ALIGNMENTS

RESULT 1
US-09-879-959-1
; Sequence 1, Application US/09879959
; Patent No. US20020160489A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H
; APPLICANT: Kumari, Kshama
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: STREPTOCOCCUS EQUISIMILIS HYALURONAN SYNTHASE GENE AND EXPRESS
; TITLE OF INVENTION: IN BACILLUS SUBTILIS
; FILE REFERENCE: 3554.049
; CURRENT APPLICATION NUMBER: US/09/879,959
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 09/469,200
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1254
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-879-959-1

Alignment Scores:

Pred. No.: 1.97e-238 Length: 1254
 Score: 2145.00 Matches: 417
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-09-469-200D-2 (1-417) x US-09-879-959-1 (1-1254)

QY 1 MetArgThrLeuLysAsnLeuLerThrValValAlaPheSerIlePheTrpValLeuLeu 20
 DB 1 ATGAGAACATTAAAAACCTCATAACTGTGTGGCCTTAGTATTTTGGGTACTGTG 60
 QY 21 IleTyrValAsnValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeu 40
 DB 61 ATTACGTCATCTTATCTCTTGGTGGCTAAGAAGCTTGTCATTTATGGCTTTTG 120
 QY 41 LeuIleAlaTyrLeuLeuValLysMetSerLeuSerPhePheTyrLysProPheLysGly 60
 DB 121 CTGATAGCTTACCTATTAGTCAAAATGCTTATGCTTTTACAAGCCATTAAAGGGA 180
 QY 61 ArgAlaGlyGlnTyrLysValAlaAlaIleProSerTyrAsnGluAspAlaGluSer 80
 DB 181 AGGCTGGGCAATATAAGTTGCGACCTATTATCCCTCTTATAACGAAGATCGTAGTCA 240
 QY 81 LeuLeuGluThrLeuLysSerValGlnGlnThrTyrProLeuAlaGluIleTyrVal 100
 DB 241 TTGCTAGAGACCTTAAAAAGTGTTCAGCAGCAAACTATCCCTAGCAGAAATTTATGTT 300
 QY 101 ValAspAspGlySerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgAsp 120
 DB 301 GTTGACATGGAGTGGCTGATGACAGAGTATTAAAGCGCATTAAGCGACTATGCGGTGAC 360
 QY 121 ThrGlyAspLeuSerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArg 140
 DB 361 ACTGGTCACTATCAGCAATGCTATGTTTCATCGCTCAGAGAAATCAAGAAACGGT 420
 QY 141 HisAlaGlnAlaTrpAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSer 160
 DB 421 CATGCAGAGCCTGGGCGCTTGAAGATCAGACGCTGATGCTTTTGGCCGTGACTCA 480
 QY 161 AspThrTyrIleTyrProAspAlaLeuGluLeuLeuLysThrPheAsnAspProThr 180
 DB 481 GATACCTTATATCCCTGATGCTTATAGAGGAGTGTGTTAAAAACCTTTAATGACCAACT 540
 QY 181 ValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArg 200
 DB 541 GTTTTTCCTGCGAGGGTCACTTAAATGTCAGAAATAGACAAACCAATCTCTTAACACGC 600
 QY 201 LeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThr 220
 DB 601 TTGACAGATATCGCTATGATAATGCTTTTGGCGTTGAACGAGCTGCCAATCCGTTACA 660
 QY 221 GlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgArgGluValValPro 240
 DB 661 GGTAAATATCCTGTGTGCTAGGTCGCTTACGCTTACAGACCGAGGTGTGTTCTCT 720
 QY 241 AsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAsp 260
 DB 721 AACATAGATAGATACATCAACACAGACCTCTCGGTATTCCTGATAGTATGGTGTGAC 780
 QY 261 ArgCysLeuThrAsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLys 280
 DB 781 AGGTGCTGTGACCACTATGCAACTGATTAGGAAGAGCTGTTTATCAATCCACTGCTAAA 840
 QY 281 CysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrpAsn 300
 DB 841 TGTATTACAGATGTTCCGTGACAGATGCTACTTACTTGAAGCAGCAAAACCGGTGGAAC 900
 QY 301 LysSerPhePheArgGluSerIleIleSerValLysLysIleMetAsnAsnProPheVal 320
 DB 901 AAGTCTCTTCTTAGACAGTCCATTTTCTGTTAAGAAATCATGAACAATCCTTTTGTGA 960

RESULT 2

US-10-011-768B-1

; Sequence 1, Application US/10011768B
 ; Publication No. US20030073221A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Weigel, Paul H.
 ; APPLICANT: DeAngelis, Paul
 ; APPLICANT: Kumari, Kshama
 ; TITLE OF INVENTION: Hyaluronan Synthase Gene and Uses Thereof
 ; FILE REFERENCE: 3554.011
 ; CURRENT APPLICATION NUMBER: US/10/011,768B
 ; CURRENT FILING DATE: 2001-12-11
 ; PRIOR APPLICATION NUMBER: US 09/178,851
 ; PRIOR FILING DATE: 1998-10-26
 ; PRIOR APPLICATION NUMBER: US 60/064,435
 ; PRIOR FILING DATE: 1997-10-31
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 1254
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 US-10-011-768B-1

Alignment Scores:

Pred. No.: 1.97e-238 Length: 1254
 Score: 2145.00 Matches: 417
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 14 Gaps: 0

US-09-469-200D-2 (1-417) x US-10-011-768B-1 (1-1254)

QY 1 MetArgThrLeuLysAsnLeuLerThrValValAlaPheSerIlePheTrpValLeuLeu 20
 DB 1 ATGAGAACATTAAAAACCTCATAACTGTGTGGCCTTAGTATTTTGGGTACTGTG 60
 QY 21 IleTyrValAsnValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeu 40
 DB 61 ATTACGTCATCTTATCTCTTGGTGGCTAAGAAGCTTGTCATTTATGGCTTTTG 120
 QY 41 LeuIleAlaTyrLeuLeuValLysMetSerLeuSerPhePheTyrLysProPheLysGly 60
 DB 121 CTGATAGCTTACCTATTAGTCAAAATGCTTATGCTTTTACAAGCCATTAAAGGGA 180
 QY 61 ArgAlaGlyGlnTyrLysValAlaAlaIleProSerTyrAsnGluAspAlaGluSer 80
 DB 181 AGGCTGGGCAATATAAGTTGCGACCTATTATCCCTCTTATAACGAAGATCGTAGTCA 240
 QY 81 LeuLeuGluThrLeuLysSerValGlnGlnThrTyrProLeuAlaGluIleTyrVal 100

```
Db 241 TTGCTAGACAGCTTAAAAAGTTTCAGCAGCAAAACCTATCCCTAGCAGAAATTTATGTT 300
QY 101 ValAspGlySerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgAsp 120
Db 301 GTTGACGATGAAGTGTGATGAGACAGGTATTAAAGCGATTGAACACTATGTGCGTGAC 360
QY 121 ThrGlyAspLeuSerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArg 140
Db 361 ACHGGTGACCTATCAGCAATGTCATTGTTTCATCGGTGAGAGAAAATCAAGAAAGCGT 420
QY 141 HisAlaGlnAlaTrpAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSer 160
Db 421 CATGCCAGAGCGCTGGCGCTTTGAAGATCAGACGCTGATGCTCTTTTGACCGTTGACTCA 480
QY 161 AspThrTyrIleTyrProAspAlaLeuGluGluLeuLysThrPheAsnAspProThr 180
Db 481 GATACCTATATCTACCTGATGCTTGAAGATCAGACGCTGATGCTCTTTTGACCGTTGACTCA 540
QY 181 ValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArg 200
Db 541 GTTTTTCGTGCGCGGTCACTTAATGTGAGAAAATAGCAAAACCAATCTCTTAACACGC 600
QY 201 LeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaGlnSerValThr 220
Db 601 TTGACAGATATTCGCTATGATATGCTTTTGGCGTTGAACGAGCTGCCCAATCCGTTACA 660
QY 221 GlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgArgGluValValPro 240
Db 661 GGTAAATATCTTGTGCTCAGTCCGCTTAGCGTTTACAGACGCGAGGTGGTGTGCTCT 720
QY 241 AsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAsp 260
Db 721 AACATAGATAGATACATCAACACGACCTTCTCGGTGATTCCTGTAAGTATTTGGTGATGAC 780
QY 261 ArgCysLeuThrAsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLys 280
Db 781 AGTGCTGTGACCACTATGCACTGATTTAGGAAGACTGTTTATCAACCACTGCTATAA 840
QY 281 CysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrpAsn 300
Db 841 TGTATTACAGATGTTCTGACAGATGCTACTTACTTGTAAAGAAAATCATGCAACATCTTGTGA 960
QY 301 LysSerPhePheArgGluSerIleLysSerValLysLysIleMetAsnAsnProPheVal 320
Db 901 AAGTCTCTCTTTAGAGAGTCCATTTATCTGTTAAGAAAATCATGCAACATCTTGTGA 960
QY 321 AlaLeuTrpThrIleLeuGluValSerMetPheMetMetLeuValTyrSerValValAsp 340
Db 961 GCCCTATGGACCATCTAGGTGCTATGTTTATGATGCTGTTTATCTGTTGTTGAT 1020
QY 341 PhePheValGlyAsnValArgGluPheAspTrpLeuArgValLeuAlaPheLeuValIle 360
Db 1021 TCTTTTGTAGCAATGTCAGAGAAATTCATTGCTCAGGCTTTCAGGCTTTCGTGAT 1080
QY 361 IlePheIleValAlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPhe 380
Db 1081 ATCTTCATGTTGCGCTGTGCGGAACATTCATTACATGCTTTAAGCACCGCTGCTCTTC 1140
QY 381 LeuLeuSerProPheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyr 400
Db 1141 TTGTTATCCGCTTTATGGGTGCTGATTTGTTGCTTACAGCCCTTGAATATAT 1200
QY 401 SerLeuPheThrIleArgAsnAlaAspTrpGlyThrArgLysLysLeuLeu 417
Db 1201 TCTCTTTTACTATTAGAAATGCTGACTGGGGAACAGTAAAAAATATTA 1251
```

RESULT 3

US-10-011-771B-1

; Sequence 1, Application US/10011771B

; Publication No. US20030082780A1

; GENERAL INFORMATION:

; APPLICANT: Weigel, Paul H.

; APPLICANT: DeAngelis, Paul

```
; APPLICANT: Kumari, Kshama
; TITLE OF INVENTION: Hyaluronan Synthase Gene and Uses Thereof
; FILE REFERENCE: 3554.011
; CURRENT APPLICATION NUMBER: US/10/011.771B
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: US 09/178,851
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: US 60/064,435
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO. 1
; LENGTH: 1254
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-10-011-771B-1
```

Alignment Scores:

```
Pred. No.: 1,97e-238 Length: 1254
Score: 2145.00 Matches: 417
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
```

US-09-469-200D-2 (1-417) x US-10-011-771B-1 (1-1254)

```
QY 1 MetArgThrLeuLysAsnLeuIleThrValValAlaPheSerIlePheTrpValLeuLeu 20
Db 1 ATGAGAACAATTAATAAAACCTCATAACTGTGTGGCCTTAGTATTTTGGGTACTGTTG 60
QY 21 IleTyrValAsnValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeu 40
Db 61 ATTACGCTCAATGTTTATCTCTTTGGTGTCTAAAGAAAGCTTGCAATTTATGGCTTTTG 120
QY 41 LeuIleAlaTyrLeuLeuValLysMetSerLeuSerPhePheTyrLysProPheLysGly 60
Db 121 CTGATACATTACCTATTAGTCAAAATGCTCTTATCCCTTTTACAGCAATTTAAGGGA 180
QY 61 ArgAlaGlyGlnTyrLysValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSer 80
Db 181 AGGCTGGCAATATAAGGTTGCGAGCCATTATTCCTCTTATAACGAAGATGCTGAGTCA 240
QY 81 LeuLeuGluThrLeuLysSerValGlnGlnThrTyrProLeuAlaGluIleTyrVal 100
Db 241 TTGCTAGAGACCTTAAAAGTGTTCAGAGCAAAACCTATCCCTAGCAAAATTTATGTT 300
QY 101 ValAspAspLysSerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgAsp 120
Db 301 GTTGACGATGGAAGTGTGATGAGACAGGTATTAAAGCGATTGAACACTATGTGCGTGAC 360
QY 121 ThrGlyAspLeuSerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArg 140
Db 361 ACTGGTGACCTATCAAGCAATGTCATTGTTTCATCGGTGAGAGAAAATCAAGAAAGCGT 420
QY 141 HisAlaGlnAlaTrpAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSer 160
Db 421 CATGCCAGAGCGCTGGCGCTTTGAAGATCAGACGCTGATGCTCTTTTGACCGTTGACTCA 480
QY 161 AspThrTyrIleTyrProAspAlaLeuGluGluLeuLysThrPheAsnAspProThr 180
Db 481 GATACCTATATCTACCTGATGCTTTAGAGAGTGTGTTAAAACCTTTAATGACCAACT 540
QY 181 ValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArg 200
Db 541 GTTTTTCGTGCGCGGTCACTTAATGTGAGAAAATAGCAAAACCAATCTCTTAACACGC 600
QY 201 LeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaGlnSerValThr 220
Db 601 TTGACAGATATTCGCTATGATATGCTTTTGGCGTTGAACGAGCTGCCCAATCCGTTACA 660
QY 221 GlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgArgGluValValPro 240
```

Db 661 GGTAATATCCCTTTGCTCAGGTCCGCTTAGCCTTTACAGACGCGAGGTGGTTGCTT 720
 QY 241 AsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAsp 260
 Db 721 ACATAGATAGATACATCAACAGACCTTCTGGGTATTCCTGTAAGTATTTGGTGATGAC 780
 QY 261 ArgCysLeuThrAsnTyrAlaThrAspLeuGlyIleThrValTyrGlnSerThrAlaLys 280
 Db 781 AGTGCTTGACCAACTATCAACTGATTTAGGAAGACCTGTTATCATCCACTGCTAAA 840
 QY 281 CysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrpAsn 300
 Db 841 TGTATACAGATGTTCTCACAAGATGCTACTTACTTGAAGACGCAAAACCGCTGGAAC 900
 QY 301 LysSerPhePheArgGluSerIleLeuValLysLysIleMetAsnAsnProPheVal 320
 Db 901 AGTCTCTTTAGAGAGTCCATATTTCTGTTAAGAAATCATGAACATCCCTTTGTA 960
 QY 321 AlaLeuTrpThrIleLeuGluValSerMetPheMetMetLeuValTyrSerValValAsp 340
 Db 961 GCCCTATGGACCATACTTGGGTGCTATGTTATGATGCTGTTTATCTGTTGTTGAT 1020
 QY 341 PhePheValGlyAsnValArgLysPheAspTrpLeuArgValLeuAlaPheLeuValIle 360
 Db 1021 TTTCTTTAGGCAATGTCAGAAATTTGATTTGCTCAGGTTTATGCTTTCTGTTGAT 1080
 QY 361 IlePheIleValAlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPhe 380
 Db 1081 ATCTTCATTTGCTGCTGTGCGGAACATTCATACATGCTTAAGACCGCTGCTTTC 1140
 QY 381 LeuLeuSerProPheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyr 400
 Db 1141 TTGTTATCTCGCTTTATGGGTGCTGCTATTTGTTGCTACAGCCCTTGAATATAT 1200
 QY 401 SerLeuPheThrIleArgAsnAlaAspTrpGlyThrArgLysLysLeuLeu 417
 Db 1201 TCTCTTTTACTATTAGAATCTGACTGGGAACACGTAAAAAATTATTA 1251

RESULT 4

US-10-172-527-1
 ; Sequence 1, Application US/10172527
 ; Publication No. US20030092118A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Weigel, Paul H
 ; APPLICANT: Kumari, Kshama
 ; APPLICANT: DeAngelis, Paul
 ; TITLE OF INVENTION: HYALURONAN SYNTHASE GENES AND EXPRESSION THEREOF IN BACILLUS SUBT
 ; FILE REFERENCE: 3554.048
 ; CURRENT APPLICATION NUMBER: US/10/172,527
 ; PRIORITY FILING DATE: 2002-06-13
 ; PRIOR APPLICATION NUMBER: 60/297,788
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: 60/297,744
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: 09/469,200
 ; PRIOR FILING DATE: 1999-12-21
 ; PRIOR APPLICATION NUMBER: 09/178,851
 ; PRIOR FILING DATE: 1998-10-26
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 1254
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 US-10-172-527-1

Alignment Scores:
 Pred. No.: 1,97e-238 Length: 1254
 Score: 2145.00 Matches: 417
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 14 Gaps: 0

US-09-469-200D-2 (1-417) x US-10-172-527-1 (1-1254)

QY 1 MetArgThrLeuLysAsnLeuIleThrValValAlaPheSerIlePheIleValLeuLeu 20
 Db 1 ATGAGAACATTAATAAAACCTCATACCTGTTGGCCCTTTAGTATTTTTGGGTACTGTTG 60
 QY 21 IleTyrValAsnValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeu 40
 Db 61 ATTAGTCAATGTTTATCTCTTTGGTGTAAAGGAAGCTTGTCAATTTATGCTGTTTGG 120
 QY 41 LeuIleAlaTyrLeuLeuValLysMetSerLeuSerPhePheTyrLysProPheLysGly 60
 Db 121 CTGATAGCTTACCTTATAGTCAAAATGCTCTTATCTCTTTTACAAGCCATTAAGGA 180
 QY 61 ArgAlaGlyGlnTyrLysValAlaIleLeuProSerTyrAsnGluAspAlaGluSer 80
 Db 181 AGGCTGGGCAATATAAGGTTGCAGCAATATTCCTCTTATAACGAAGATGCTGAGTCA 240
 QY 81 LeuLeuGluThrLeuLysSerValGlnGlnGlnThrTyrProLeuAlaGluIleTyrVal 100
 Db 241 TTGCTAGAGACCTTAATAAAGTGTTCAGCAGCAAAACCTATCCCTACAGCAAAATTTATG 300
 QY 101 ValAspAspGlySerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgAsp 120
 Db 301 GTTGAGATGGAAGTGTGATGAGCAGGTATTAAGCGCATTAAGACATGATGCGTGAC 360
 QY 121 ThrGlyAspLeuSerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArg 140
 Db 361 ACTGGTGACCTATCAAGCAATGTCTATGTCGTCAGAGAAAAATCAAGGAAGCCT 420
 QY 141 HisAlaGluAlaIleThrAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSer 160
 Db 421 CATGCAAGCCCTGGGCTTTGAAGATCAGACGCTGATGCTCTTTTACCGCTGACTCA 480
 QY 161 AspThrTyrIleTyrProAspAlaLeuGluLeuLysThrPheAsnAspProThr 180
 Db 481 GATACCTATATCTACCTGATGCTTTAGAGGAGTTGTTAAAAACCTTTATGACCAACT 540
 QY 181 ValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArg 200
 Db 541 GTTTTTGCTGGCAGCGGTACCTTAATGTCAAGAAATAGACAAACCAATCTCTTAACACGC 600
 QY 201 LeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThr 220
 Db 601 TTGACAGATTCGCTGATGATGATGCTTTGGCTTGAAGAGCTGCCCAATCCGTTACA 660
 QY 221 GlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgGluValValPro 240
 Db 661 GGTAATATCTTTGCTCAGGTCGCTTAGCGTTTACAGACGCGAGGTGGTTGTTCT 720
 QY 241 AsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAsp 260
 Db 721 ACATAGATAGATACATCAACAGACCTTCCTGGGTATTCCTGTAAGTATTTGGTGATGAC 780
 QY 261 ArgCysLeuThrAsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLys 280
 Db 781 AGGTGCTTGACCAACTATGCACTGATTTAGGAAGACCTGTTTATCATCCACTGCTAAA 840
 QY 281 CysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrpAsn 300
 Db 841 TGTATACAGATGTTCTCACAAGATGCTACTTACTTACTTGAAGCAGCAAAACCGCTGGAAC 900
 QY 301 LysSerPhePheArgGluSerIleLeuValLysLysIleMetAsnAsnProPheVal 320
 Db 901 AGTCTCTTTAGAGAGTCCATATTTCTGTTAAGAAATCATGAACATCCCTTTGTA 960
 QY 321 AlaLeuTrpThrIleLeuGluValSerMetPheMetMetLeuValTyrSerValValAsp 340
 Db 961 GCCCTATGGACCATACTTGGGTGCTATGTTATGATGCTGTTTATCTGTTGTTGAT 1020
 QY 341 PhePheValGlyAsnValArgLysPheAspTrpLeuArgValLeuAlaPheLeuValIle 360

Db 1021 TCTCTGTAGGCAATGTGAGAAATTTGATTGGCTCAGGGTTTTCAGCGTTTTCGGTGATT 1080
 QY 361 IlePheIleValAlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPhe 380
 Db 1081 ATCTTCATTGTTCCTGTGTCGACATTCATTACATGCTTAAGCACCCGCTGCTCTTC 1140
 QY 381 LeuLeuSerProPheTyrGlyValLeuHisLeuPheValLeuGluProLeuLysLeuTyr 400
 Db 1141 TTTGTTATCCGTTTATGAGGGTGTGCTGATTTGTTGCTACAGCCCTTGAATATATAT 1200
 QY 401 SerLeuPheThrIleArgAsnAlaAspTrpGlyThrArgLysLysLeuLeu 417
 Db 1201 TCTCTTTTACTATTAGAAATGCTGACTGGGGAACAGTAAAAAATTATTA 1251

RESULT 5

US-10-326-185-1
 ; Sequence 1, Application US/10326185
 ; Publication No. US20030175902A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sloma, Alan
 ; APPLICANT: Behr, Regine
 ; APPLICANT: Widner, William
 ; APPLICANT: Tang, Maria
 ; APPLICANT: Sternberg, David
 ; APPLICANT: Brown, Stephen
 ; TITLE OF INVENTION: Methods for Producing Hyaluronan In a Recombinant Host Cell
 ; FILE REFERENCE: 10241.200-US
 ; CURRENT APPLICATION NUMBER: US/10/326,185
 ; CURRENT FILING DATE: 2002-12-20
 ; PRIOR APPLICATION NUMBER: US 60/342,644
 ; PRIOR FILING DATE: 2001-12-21
 ; NUMBER OF SEQ ID NOS: 108
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 1251
 ; TYPE: DNA
 ; ORGANISM: Streptococcus equisimilis
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1251)
 ; OTHER INFORMATION:
 US-10-326-185-1

Alignment Scores:

Pred. No.: 1.27e-237 Length: 1251
 Score: 2138.00 Matches: 416
 Percent Similarity: 99.76% Conservative: 0
 Best Local Similarity: 99.76% Mismatches: 1
 Query Match: 99.67% Indels: 0
 DB: 12 Gaps: 0

US-09-469-200d-2 (1-417) x US-10-326-185-1 (1-1251)

QY 1 MetArgThrLeuLysAsnLeuIleThrValValAlaPheSerIlePheTrpValLeuLeu 20
 Db 1 ATGAGACATTAATAAAACCTCATCTACTGTGGCCCTTAGTATTTTGGGTACTGTG 60
 QY 21 IleTyrValAsnValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeu 40
 Db 61 ATTACGTCAATGTTATCTCTTGTGTCTAAAGGAAGCTTGTCAATTATGCTTTTG 120
 QY 41 LeuIleAlaTyrLeuLeuValLysMetSerLeuSerPheTyrTyrProPheLysGly 60
 Db 121 CTGATAGCTTACTATTAGTCAAAATGTCTTATCCCTTTTATACAGCCATTTAAGGA 180
 QY 61 ArgAlaGlyGlnTyrLysValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSer 80
 Db 181 AGGCTGGGCAATATAGGTTCAGCCATTATCCCTCTTATACGAAGATGCTGAGTCA 240
 QY 81 LeuLeuGluThrLeuLysSerValGlnGlnThrTyrProLeuAlaGluIleTyrVal 100
 Db 241 TTGCTAGACCTTAAAAAGTTCAGCAGCAACCTATCCCTAGCAGAAATTTATGTT 300

QY 101 ValAspAspGlySerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgAsp 120
 Db 301 GTTGACCATGGAAGTCTCTGATGACAGAGTATTAAGCGCATTAAGACTATGTGGTGAC 360
 QY 121 ThrGlyAspLeuSerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArg 140
 Db 361 ACTGGTGACCTATCAAGCAATGTCTATGTTTCACCGGTGAGAAAATCARGAAGCGT 420
 QY 141 HisAlaGlnAlaTrpAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSer 160
 Db 421 CATGCACAGGCGTGGCCCTTTGAAAGATCAGACGCTGATGCTCTTTTGACCGTTCAC 480
 QY 161 AspThrTyrIleTyrProAspAlaLeuGluLeuLeuLysThrPheAsnAspProThr 180
 Db 481 GATCTTATATCTACCTGATGCTTTAGAGGAGTGTATAAAACCTTTAATGACCAACT 540
 QY 181 ValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArg 200
 Db 541 GTTTTGTGCGAGCGGTCACTTAATGTCAGAAATAGACAAACCAATCTCTTAACAGC 600
 QY 201 LeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThr 220
 Db 601 TTGACAGATATTCGCTATGATAATGCTTTTGGCGTTGAACGAGTGCCTCAATCCCTTACA 660
 QY 221 GlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgArgGluValValPro 240
 Db 661 GGTAATATCTCTGTTGCTCAGGCGCCCTTAGCGTTTACAGCGAGGTGGTGTTCCT 720
 QY 241 AsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAsp 260
 Db 721 AACATAGATAGATACATCAACGACGCTTCCTGGGTATTCCTGTAAAGTATCGGTGATGAC 780
 QY 261 ArgCysLeuThrAsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLys 280
 Db 781 AGGTGCTGTGACCAACTATGCAACTGATTTAGGAAAGACTGTTTATCAATCCACTGCTAAA 840
 QY 281 CysIleThrAspValProAspLysMetSerTyrTyrLeuLysGlnGlnAsnArgTrpAsn 300
 Db 841 TGTATACAGATGCTCTGACAGATGCTCTACTTACTTGAAGCAGCAAAACCGCTGGAAC 900
 QY 301 LysSerPhePheArgGluSerIleIleSerValLysLysIleMetAsnAsnProPheVal 320
 Db 901 AAGTCCTCTTTAGAGAGTCCATTATTCGTTAAGAAAATCATGAACAATCCTTTTGTGA 960
 QY 321 AlaLeuTrpThrIleLeuGluValSerMetPheMetMetLeuValTyrSerValValAsp 340
 Db 961 GCCCTATGGACCATCTATGAGTGTCTATGTTTATGATGCTTGTATTCTCTGTGGTGGAT 1020
 QY 341 PhePheValGlyAsnValArgGluPheAspTrpLeuArgValLeuAlaPheLeuValIle 360
 Db 1021 TTTCTTTGTAGACAAATGTCAGAGAATTTGATGGCTCAGGGTTTGGCCTTCTGGTGATT 1080
 QY 361 IlePheIleValAlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPhe 380
 Db 1081 APTCTTCATTTGCTCTTTTGTGTAATATATCACTATATGCTTAAGCACCCGCTGCTCTTC 1140
 QY 381 LeuLeuSerProPheTyrGlyValIleHisLeuPheValLeuGlnProLeuLysLeuTyr 400
 Db 1141 TTTGTTATCCGTTTATGAGGGTACTGCAATTTGTTGCTCAGACCCCTTGAATTTGAT 1200
 QY 401 SerLeuPheThrIleArgAsnAlaAspTrpGlyThrArgLysLysLeuLeu 417
 Db 1201 TCTCTTTTACTATTAGAAATGCTGACTGGGGAACAGTAAAAAATTATTA 1251

RESULT 6

US-10-326-185-108
 ; Sequence 108, Application US/10326185
 ; Publication No. US20030175902A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sloma, Alan
 ; APPLICANT: Behr, Regine
 ; APPLICANT: Widner, William
 ; APPLICANT: Tang, Maria

```

; APPLICANT: Sternberg, David
; APPLICANT: Brown, Stephen
; TITLE OF INVENTION: Methods for Producing Hyaluronan In a Recombinant Host Cell
; FILE REFERENCE: 10241.200-US
; CURRENT APPLICATION NUMBER: US/10/326,185
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/342,644
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 108
; LENGTH: 5158
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-10-326-185-108

Alignment Scores:
Pred. No.:      8,48e-218      Length:      5158
Score:          1976.00        Matches:      383
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:     92.12%       Indels: 0
DB:              12          Gaps: 0

US-09-469-200D-2 (1-417) x US-10-326-185-108 (1-5158)

QY 35 SerileTyrGlyPheLeuLeuLeuAlaTyrLeuLeuValLysMetSerLeuSerPhePhe 54
DB 1 TCATTTATGCTTTTCTGCTATGATACCTTACCTATTAGTCAAAATGTCCTATCCCTTTT 60
QY 55 TyrLysProPheLysGlyArgAlaGlyGlnTyrLysValAlaAlaLeuLeuProSerTyr 74
DB 61 TACAAGCCATTTAAGGAAGGCTGGCAATATAGGTTCAGCCATTTATCCCTCTAT 120
QY 75 AsnGluAspAlaGluSerLeuLeuGluThrLeuLysSerValGlnGlnGlnThrTyrPro 94
DB 121 AACCAAGATCTGAGTCAATCTGAGACCTTAAAGATGTTCAAGCAAAACCTATCCC 180
QY 95 LeuAlaGluLeuTyrValValAspGlySerAlaAspGluThrGlyLeuLysArgTle 114
DB 181 CTAGCAGAAATTTATGTTGTTGACGATGGAGTCTGATGACAGCATTTAAGCGCAT 240
QY 115 GluAspTyrValArgAspThrGlyAspLeuSerSerAsnValLeuValHisArgSerGlu 134
DB 241 CAAGACATATGCTGTCACACTGGTGCCTATCAAGCAATGTCATTCTCATCGGTACAG 300
QY 135 LysAsnGlnGlyArgHisAlaGlnAlaTyrAlaPheGluArgSerAspAlaAspVal 154
DB 301 AAAATCAAGAAAGGCTCATGACAGGCTGGCTTTGAAGATCAGACGCTGATGTC 360
QY 155 PheLeuThrValAspSerAspThrTyrTyrProAspAlaLeuGluLeuLeuLys 174
DB 361 TTTTGTGACCGTTCAGTCAATATATATCTACCTGATGCTTTAGAGGAGTTGTAAAA 420
QY 175 ThrPheAsnAspProThrValPheAlaAlaThrGlyHisLeuAsnValArgAsnGln 194
DB 421 ACCTTTATGACCAACTGTTTTTGTGCGACGCGTCACTTAAATGCTCAGAAATAGACAA 480
QY 195 ThrAsnLeuLeuThrArgLeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArg 214
DB 481 ACCAATCTCTACACGCTTCACAGATATTCCTGATGATATGCTTTTGGCGTTGAACGA 540
QY 215 AlaAlaGlnSerValThrGlyAsnIleLeuValCysSerGlyProLeuSerValTyrArg 234
DB 541 GCTGCCCAATCCGTTACAGGTAATATCTTTGCTCAGGTCGCTTACGCTTACGTTACAGA 600
QY 235 ArgGluValValProAsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIlePro 254
DB 601 CCGGAGGTGTTGCTTCAACATAGATAGATACATCAACACAGACCTTCTCGGTATTCCT 660
QY 255 ValSerIleGlyAspAspArgCysLeuThrAsnTyrAlaThrAspLeuGlyIysThrVal 274
DB 661 GTAAGTATTGGTGATGACAGGTGCTTGACCAACTATGCACTGATTTAGGAAGACTGTT 720

```

```

QY 275 TyrGlnSerThrAlaLysCysIleThrAspValProAspLysMetSerThrTyrLeuLys 294
DB 721 TATCAATCCACTGCTAAATGATATTACAGATGTTCTCTGACAAGATGCTCTACTTGAAG 780
QY 295 GlnGlnAsnArgTyrAsnLysSerPhePheArgGluSerIleIleSerValLysIle 314
DB 781 CAGCAAAACCGCTGGAACAGTCTTTTAGAGAGTCCATTAATTTCTGTAAAGAAATC 840
QY 315 MetAsnAsnProPheValAlaLeuThrPheIleLeuGluValSerMetPheMetMetLeu 334
DB 841 ATGACAAATCTTTTGTAGCCCTATGGACCATCTGAGGTGCTATGTTTATGATGCTT 900
QY 335 ValTyrSerValValAspPhePheValGlyAsnValArgGluPheAspTyrLeuArgVal 354
DB 901 GTTTATTCTGTGGTGGATTTCTTTAGGCAATGTCAGAGAAATTTGATGGCTCAGGGTT 960
QY 355 LeuAlaPheLeuValIleIlePheIleValAlaLeuCysArgAsnIleHisTyrMetLeu 374
DB 961 TTAGCCCTTTCTGGTGATTAICTTCAITGTTGCCCTGTGCGGACATTCATTACATGCTT 1020
QY 375 LysHisProLeuSerPheLeuLeuSerProPheTyrGlyValLeuHisLeuPheValLeu 394
DB 1021 AAGCACCCGCTGCTCTCTTGTATCTCCGTTTATGGGTGCTGCATTTGTTTGTCTTA 1080
QY 395 GlnProLeuLysLeuTyrSerLeuPheThrIleArgAsnAlaAspTyrGlyThrArgLys 414
DB 1081 CAGCCCTTGAATATATTTCTCTTTTACTATTAGAAATGCTGACTGGGAACACGTAAA 1140
QY 415 LysLeuLeu 417
DB 1141 AATTTATTA 1149

RESULT 7
US-10-326-185-102
; Sequence 102, Application US/10326185
; Publication No. US20030175902A1
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; APPLICANT: Behr, Regine
; APPLICANT: Widner, William
; APPLICANT: Tang, Maria
; APPLICANT: Sternberg, David
; APPLICANT: Brown, Stephen
; TITLE OF INVENTION: Methods for Producing Hyaluronan In a Recombinant Host Cell
; FILE REFERENCE: 10241.200-US
; CURRENT APPLICATION NUMBER: US/10/326,185
; PRIOR FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/342,644
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 102
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Streptococcus uberis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1251)
; OTHER INFORMATION:
US-10-326-185-102

Alignment Scores:
Pred. No.:      2,67e-186      Length:      1251
Score:          1695.50        Matches:      312
Percent Similarity: 87.47%    Conservative: 51
Best Local Similarity: 75.18% Mismatches: 51
Query Match:     79.04%       Indels: 1
DB:              12          Gaps: 1

US-09-469-200D-2 (1-417) x US-10-326-185-102 (1-1251)

QY 1 MetArgThrLeuLysAsnLeuIleThrValValAlaPheSerIlePheTrpValLeuLeu 20

```

```

Db 1 ATGGAACAACTAAACAACTCAATACATTTATGACCTTTATTTCCCTGGCCCAATAT 60
QY 21 IletyValAsnValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeu 40
Db 61 ATTGGGCTTAATGTTTGTATTTGGAACCTAAAGGAAGTCTAACAGGTGTATGGATATT 120
QY 41 LeuLeuAlaTyrLeuLeuValLysMetSerLeuSerPhePheTyrLysPropheLysGly 60
Db 121 CTATTAACTATTGTCGATGATAAAATGGGATTATCTTTTATTCGTCCTATAAAGA 180
QY 61 ArgAlaGlyInTyrLysValAlaAlaIleProSerTyrAsnGluAspAlaGluSer 80
Db 181 AGTGTAGTCAATATAGTAGTACGACTATTATCCATCTTATATAGGATGGTGTGGT 240
QY 81 LeuLeuGluThrLeuLysSerValGlnGlnThrTyrProLeuAlaGluIleTyrVal 100
Db 241 TTACTAGAACTCTTAAGAGTGTTCAAAACCAACATATCCAAATGGCAGAAATTTTCGTA 300
QY 101 ValAspAspGlySerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgAsp 120
Db 301 ATTGAGATGGTCTAGTATATAAACAGGTATAAATGGTCGAGACTATGGAAGTTA 360
QY 121 ThrGlyAspLeuSerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArg 140
Db 361 AATGGC---TTTGGAGACCAAGTTATCGTTCATCAGATCCCTGAAATGTTGCTAAAGA 417
QY 141 HisAlaGlnAlaTyrAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSer 160
Db 418 CAGCTCAGGCTGGGATTTGAAAGGCTGATGCTGATGTTTCTTAAACAGTGGATCA 477
QY 161 AspThrTyrIleTyrProAspAlaLeuGluLeuLeuLysThrPheAsnAspProThr 180
Db 478 GATACCTACATCTATCTGATGCTCTTGAAGATTTATTAAGACATTTATGATCCAGAG 537
QY 181 ValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArg 200
Db 538 GTCTAGCTGCACTGCTCATTTAAATGCAAGAAATAGACAACTTAATCTCTTAACATA 597
QY 201 LeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThr 220
Db 598 CTGCTGATATCTGTTACGATATGCAATTTGGTGAACGGTGGCTCAGTCTGTACG 657
QY 221 GlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgArgGluValValPro 240
Db 658 GGAATATTGTTGTTCCGACCTTTAAGTATTATAGAGCTTCGCTGGTATCCA 717
QY 241 AsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAsp 260
Db 718 AATCTGTAACGCTATACCTACAAACATTTCTTGGTGCCTGTAAGCATGGGATGAC 777
QY 261 ArgCysLeuThrAsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLys 280
Db 778 CGTGTGTTGCAAAATATGCACTGATTTGGGAAAACGGTTATCAGTCACTGCAAGA 837
QY 281 CysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTyrAsn 300
Db 838 TGTGATCTGACGTTCCAGATGTTTAAGGTTTTCATCAACAAACAAATCGTTGGAT 897
QY 301 LysSerPhePheArgGluSerIleIleSerValLysLysIleMetAsnAsnProPheVal 320
Db 898 AAGTCATTTTTAGGGAGTCTATATCTCTTTAAGAAGTTATTAGCCACCAAGTGT 957
QY 321 AlaLeuIleThrIleLeuGluValSerMetPheMetLeuValTyrSerValValAsp 340
Db 958 CGCTGTGAGCTATTACAGAGTTTCCATGTTTCATCATGCTAGTTTATCTCTTTAGC 1017
QY 341 PhePheValGlyAsnValArgGluPheAspTyrLeuArgValLeuAlaPheLeuValIle 360
Db 1018 TTATCTAGGAGGCTCAAGAAATTTAATCTCATAAACTGTTGCTTTTATGATTAT 1077
QY 361 IlePheIleValAlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPhe 380

```

```

Db 1078 ATTTTCATAGTAGCTCTTTGTAGAAATGTTCAATACATGTTAAGCATCCATTTGCTTTT 1137
QY 381 LeuLeuSerPropheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyr 400
Db 1138 TTATGTGCACCGTTTATATGATTGATACATCTATTCTGTTTGCACCTCTTAAGATATAT 1197
QY 401 SerLeuPheThrIleArgAsnAlaAspTyrGlyThrArgLysLys 415
Db 1198 TCGTATTACTATAGAATGCTACATGGGAACCTGTAAGAAG 1242

RESULT 8
US-10-172-527-11
; Sequence 11, Application US/10172527
; Publication No. US20030092118A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H
; APPLICANT: Kumari, Kshama
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: HYALURONAN SYNTHASE GENES AND EXPRESSION THEREOF IN BACILLUS S
; FILE REFERENCE: 3554, 048
; CURRENT APPLICATION NUMBER: US/10/172, 527
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/297,788
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 60/297,744
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 09/469,200
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 3466
; TYPE: DNA
; ORGANISM: Streptococcus uberis
US-10-172-527-11

Alignment Scores:
Pred. No.: 1,48e-185 Length: 3466
Score: 1895.50 Matches: 312
Percent Similarity: 87.47% Conservative: 51
Best Local Similarity: 75.18% Mismatches: 51
Query Match: 79.04% Indels: 1
DB: 14 Gaps: 1

US-09-469-200D-2 (1-417) x US-10-172-527-11 (1-3466)
QY 1 MetArgThrLeuLysAsnLeuIleThrValValAlaPheSerIlePheTyrValLeuLeu 20
Db 23 ATGGAACAACTAAACAACTCAATACATTTATGACCTTTATTTCCCTGGCCCAATAT 82
QY 21 IletyValAsnValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeu 40
Db 83 ATTGGGCTTAATGTTTGTATTTGGAACCTAAAGGAAGTCTAACAGGTGTATGGATATT 142
QY 41 LeuLeuAlaTyrLeuLeuValLysMetSerLeuSerPhePheTyrLysPropheLysGly 60
Db 143 CTATTAACTATTGTCGATGATAAAATGGGATTATCTTTTATTCGTCCTATAAAGA 202
QY 61 ArgAlaGlyGlnTyrLysValAlaAlaIleProSerTyrAsnGluAspAlaGluSer 80
Db 203 AGTGTAGTCAATATAGTAGTACGACTATTATCCATCTTATAATGAGATGGTGTGGT 262
QY 81 LeuLeuGluThrLeuLysSerValGlnGlnThrTyrProLeuAlaGluIleTyrVal 100
Db 263 TTACTAGAACTCTAAAGAGTGTTCAAAACCAACATATCCAAATGGCAGAAATTTTCGTA 322
QY 101 ValAspAspGlySerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgAsp 120
Db 323 ATTGACCATGGGTGCTAGTATATAAACAGGTATAAATGGTCCGAGACTATGGAAGTTA 382
QY 121 ThrGlyAspLeuSerSerAsnValIleValHisArgSerGluLysAsnGlnGlyLysArg 140

```


QY 245 TyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspArgCysLeuThr 264
 Db 730 TATAAAATCAACATCTCTAGGTTTACCTGGTTAGCATGGGGATGATGATGTTTAAACA 789
 QY 265 AsnTyrAlaThrAspLeuGlySerThrValTyrGlnSerThrAlaLysCysIleThrAsp 284
 Db 790 AATATGCTATGATTTAGGACGACGCTCTACCAATCAACAGCTAGATGATGATGAT 849
 QY 285 ValProAspLysMetSerThrTyrIleLysGlnGlnAsnArgTrpAsnLysSerPhePhe 304
 Db 850 GTACCTTTCCAAATTAAGAAATTAATTAAGCAACAAATTCGATGGAATAAATCTTTT 909
 QY 305 ArgGluSerIleLeuSerValLysLysIleMetAsnAsnProPheValAlaLeuThrPhe 324
 Db 910 AAGAAGTCTATTTCTGTAAATAAATCTTTCTAATCCCATCGTGTGCTTATGGACT 969
 QY 325 IleLeuValValSerMetPheMetMetLeuValValSerValValAspPheValGly 344
 Db 970 ATTTTCGAAGTCGTATGTTTATGATGTTGATGTCGCAATTTGGGAATCTTTTGTAT 1029
 QY 345 AsnValArgGluPheAspTrpLeuArgValLeuAlaPheLeuValIleLeuPheVal 364
 Db 1030 CAAGCTATCAATTAAGACCTTTTAAACCTTTTTCCTTTTATCCATCATCTTTATCGT 1089
 QY 365 AlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPheLeuSerPro 384
 Db 1090 GCTTTTNGCTGTAATGCTTATATATATATATATATATATATATATATATATATAT 1149
 QY 385 PheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyrSerLeuPheThr 404
 Db 1150 CTGTATGGAATATATACATGTTGTTGCTTACAGCCCTTAAACCTTTTATCTTATGCACC 1209
 QY 405 IleArgAsnAlaAspTrpGlyThrArgLysLysLeu 416
 Db 1210 ATTAATAACAGGAATGGGAACACAGTAAAAANGGTC 1245

RESULT 10

US-10-172-527-13
 ; Sequence 13, Application US/10172527
 ; Publication No. US20030092118A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Weigel, Paul H
 ; APPLICANT: Kumari, Kshama
 ; APPLICANT: DeAngelis, Paul
 ; TITLE OF INVENTION: HYALURONAN SYNTHASE GENES AND EXPRESSION THEREOF IN BACILLUS SUBT
 ; FILE REFERENCE: 3554.048
 ; CURRENT APPLICATION NUMBER: US/10/172.527
 ; CURRENT FILING DATE: 2002-06-13
 ; PRIOR APPLICATION NUMBER: 60/297,788
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: 60/297,744
 ; PRIOR FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: 09/469,200
 ; PRIOR FILING DATE: 1999-12-21
 ; PRIOR APPLICATION NUMBER: 09/178,851
 ; PRIOR FILING DATE: 1998-10-26
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 13
 ; LENGTH: 1440
 ; TYPE: DNA
 ; ORGANISM: Streptococcus pyogenes
 ; US-10-172-527-13

Alignment Scores:

Pred. No.:	1.48e-173	Length:	1440
Score:	1586.50	Matches:	298
Percent Similarity:	85.19%	Conservative:	53
Best Local Similarity:	72.33%	Mismatches:	60
Query Match:	73.96%	Indels:	1
DB:	14	Gaps:	1

US-09-469-200d-2 (1-417) x US-10-172-527-13 (1-1440)

QY 5 LysAsnLeuIleThrValValAlaPheSerIlePheTrpValLeuLeuIleTyrValAsn 24
 Db 133 AAAAAAACCTTAATGTTTATCTTATTTTGGATATCTATCTGATTATCTAAAT 192
 QY 25 ValTyrLeuPheGlyAlaLysGlySerLeuSerIleTyrGlyPheLeuLeuAlaTyr 44
 Db 193 ATGTATCTATTGGAACA---TCAACTGTAGGAATTTATGGAGTAAATTAATAACCTAT 249
 QY 45 LeuLeuValLysMetSerLeuSerPhePheTyrLysProPheLysGlyArgAlaGlyGln 64
 Db 250 CTAGTATCAACTTGGATATCTTTCTTATGAGCCATTTAAAGGAAATCCACATGAC 309
 QY 65 TyrLysValAlaAlaIleLeuProSerTyrAsnGluAspAlaGluSerLeuLeuGluThr 84
 Db 310 TATAAGTTGCTGCTGTAATCTCTTATAATGAAGTCCGAGTCAATATTAGAAACA 369
 QY 85 LeuLysSerValGlnGlnThrTyrProLeuAlaGluIleTyrValValAspAspGly 104
 Db 370 CTTAAAGTGTGTTAGCACAGACCTATCCGTTATCAGAAATTTATTTGTTGATGGG 429
 QY 105 SerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgAspThrGlyAspLeu 124
 Db 430 AGITCAACACAGATGCAATACAAATTAATGAAGTATGTAATAGAGAAAGTGGATAT 489
 QY 125 SerSerAsnValIleValHisArgSerGluLysAsnGlnGlyArgHisAlaGlnAla 144
 Db 490 TGTGAAAGCTTATGCTTCCCGTTCCCTTGTCAATTAAGGAAAAAGCCATGCTCAAGG 549
 QY 145 TrpAlaPheGluArgSerAspAlaAspValPheLeuThrValAspSerAspThrTyrIle 164
 Db 550 TGGGCATTTGAAGATCTGACGCTGACGTTTTTAAACCGTAGACTCAGATCTATATC 609
 QY 165 TyrProAspAlaLeuGluGluLeuLeuLysThrPheAsnAspProThrValPheAlaAla 184
 Db 610 TATCCAAATGCTTAGAAGAACTCTTAAAGCTTCAATGATGACACAGATTTATGTCGA 669
 QY 185 ThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArgLeuThrAspIle 204
 Db 670 ACAGGACATTTGAATGTAGAAACAGACAACTAATCTATTACGCGACTTACAGATATC 729
 QY 205 ArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThrGlyAsnIleLeu 224
 Db 730 CGTTACGATTAATGCTTTGGGGTGGAGCGTCTCTCAATCAATTAACAGTAAATTTTA 789
 QY 225 ValCysSerGlyProLeuSerValTyrArgArgGluValValValProAsnIleAspArg 244
 Db 790 GTTTGCTCAGGACCATTTAGTATTTATCGACGTGAAGTATTTCTTAACTTACAGCGC 849
 QY 245 TyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAspArgCysLeuThr 264
 Db 850 TATAAAATCAAAACATTCCTAGTTTACCTGTAGCATGGGATGATGATGTTTAAACA 909
 QY 265 AsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLysCysIleThrAsp 284
 Db 910 AATTATGCTATTGATTTAGGACGACCTGCTACCAATCAACAGCTAGATGATGATGAT 969
 QY 285 ValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrpAsnLysSerPhePhe 304
 Db 970 GTACCTTTCCAAATTAAGAAATTTAAGCAACAAATTCGATGGAATAAATCTTTT 1029
 QY 305 ArgGluSerIleLeuSerValLysLysIleMetAsnAsnProPheValAlaLeuTyrPhe 324
 Db 1030 AGAATCTATTATTTCTGTAAAAAATTTCTTCTAATCCCATCGTGTGCTTATGGACT 1089
 QY 325 IleLeuGluValSerMetPheMetMetLeuValTyrSerValValAspPhePheValGly 344
 Db 1090 ATTTTCGAAGTCTGTTTATGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 1149
 QY 345 AsnValArgGluPheAspTrpLeuArgValLeuAlaPheLeuValIleLeuPheIleVal 364
 Db 1150 CAAGCTATCAATTAGACCTTATTAACCTTTTTCCTTTTATCCATCATCTTATCGTT 1209

[illegible]

5

682	TATCCAAATGCCATTAGAGAAGTCTCTAAAAAGCTTCATCATGATGACAGAGTTTATGCTGCA	741
185	ThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArgLeuThrAspIle	204
742	ACAGACATTTGAATGCTTAGAAACAGACAAACTAATCTATTACGCGACTTACAGATATC	801
205	ArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThrGlyAsnIleLeu	224
802	CGTTACGATTAATGCCCTTTGGGGTGGAGCGTGTCTCAATCATTAACAGGTAATATTTTA	861
225	ValCysSerGlyProLeuSerValTyrArgArgGluValValProAsnIleAspArg	244
862	GTTTGCTCAGSACCATTTGAGTATTATTCAGCGTGAAGTGATTATCTTAACCTTAGAGCGC	921
245	TyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspArgCysLeuThr	264
922	TATAAAATCAAACATTCCTAGGTTTACCTGTAGTATCGGAGATCGCATGCTTTAACA	981
265	AsnTyrAlaThrAspLeuGlyLysThrValTyrGlnSerThrAlaLysCysIleThrAsp	284
982	AATTATGCTATTGATTAGGACGCAGCTGTCTACCAATCAACAGCTAGATGGATCTGAT	1041
285	ValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTyrAsnLysSerPhePhe	304
1042	GTACCTTTCCAAATTAAGATTTATTAAAGCAACAAATCGATGGAATAAATCTTTT	1101
305	ArgGluSerIleLeuSerValLysIleMetAsnAsnProPheValAlaLeuThrPthr	324
1102	AGAGATCATATTATTCTGTTAAAAAATCTTTCTTAATCCCATCGTTTGCCTTATGGACT	1161
325	IleLeuGluValSerMetPheMetLeuValTyrSerValValAspPhePheValGly	344
1162	ATTTCGGAAGTCGTTATGTTTATGATTGATGTCGCAATGGGAAATCTTTGTTTAAT	1221
345	AsnValArgGluPheAspTyrPheArgValLeuAlaPheLeuValIlePheIleVal	364
1222	CAAGCTATTCAATTAGACCTTATTAACTTTTGGCTTTTATCCCATCATCTTATCGTT	1281

1

```

QY 325 IleLeuGluValSerMetPheMetLeuValTyrSerValValAlaSpPhePheValGly 344
Db 1162 AITTTGGAAGTCGTATGCTTATCATGCTGATGTCGCAATGGGAAATCTTTTGTTTAA 1221
QY 345 AsnValArgGluPheAspTyrPheuArgValLeuAlaPheLeuValIleIlePheIleVal 364
Db 1222 CAAGCTATTCAATTAGACCTTATTAAACTTTTGGCTTTTATCCATCATCTTTATCGTT 1281
QY 365 AlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSerPheLeuLeuSerPro 384
Db 1282 GCTTTATGCTGTAATGTTTCATTATATGGTCAACAATCCTCTAGTTTGTGTTATCTCT 1341
QY 385 PheTyrGlyValLeuHisLeuPheValIleuGlnProLeuLysLeuTyrSerLeuPheThr 404
Db 1342 CTGTATGGAATATTACACTGTGTGCTTACAGCCCTTAAACTTTATCTTTATGCAC 1401
QY 405 IleArgAsnAlaAspTyrPheThrArgLysLysLeu 416
Db 1402 ATTAAATATAGGAATGGGGACACGTAAAGGTC 1437
RESULT 12
US-10-117-795-1
: Sequence 1, Application US/10117795
: Publication No. US20030104533A1
: GENERAL INFORMATION:
: APPLICANT: WEIGEL, PAUL E.
: APPLICANT: DEANGELIS, PAUL L.
: APPLICANT: PAPAKONSTANTINO, JOHN
: TITLE OF INVENTION: HYALURONATE SYNTHASE GENES AND USES THEREOF
: FILE REFERENCE: 35541.073

```


Best Local Similarity: 51.23% Mismatches: 98
 Query Match: 45.57% Indels: 17
 DB: 14 Gaps: 2

US-09-469-200D-2 (1-417) x US-10-172-527-17 (1-1200)

QY 66 LysValAlaAlaIleProSerTyrAsnGluAspAlaGluSerLeuLeuGluThrLeu 85
 DB 17 AAAGTGCAGTGTGCTCCCTCATATATGAAGTGCAGTGTATTTAATACAAAT 76
 QY 86 LysSerValGlnGlnThrTyrProLeuAlaGluIleTyrValValAspAspGlySer 105
 DB 77 AATAGCGTTTATAGTCAAGATTATCAATATCAATATCAATATTTCTTGTATGATGTAGT 136
 QY 106 AlaAsp-----GluThrGlyIleLysArgIleGluAspTyrValArgAspThr 121
 DB 137 AAGATAAATCGGCTTATGAAGTAGCACTTAAATGAGGAGGAATCTTTAGACATCAA 196
 QY 122 GlyAspLeuSerSer-----AsnVal 128
 DB 197 CGAAGAAATTCCTGCTACAACTAAGAAATTTCTGAAATATATAGTATTCCTGACTTA 256
 QY 129 IleValHisArgSerGluLysAsnGlnGlyLysArgHisAlaGlnAlaIlePalaPheGlu 148
 DB 257 ATCTGATCGTTTATAGTAAAGATTGCGGAAAGACATGCTCAATATGCGCTTTAA 316
 QY 149 ArgSerAspAlaAspValPheLeuThrValAspSerAspThrTyrIleTyrProAspAla 168
 DB 317 CGGACACACAGATGCTATGTTACCATGATGATGATGATGATGATGATGATGATGAT 376
 QY 169 LeuGluLeuLeuLeuLeuLeuPheAsnAspProThrValPheAlaIleThrGlyHisLeu 188
 DB 377 GTTAGAGATTATGAACCCCTTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 436
 QY 189 AsnValArgAsnArgGlnThrAsnLeuLeuThrArgLeuThrAspIleArgTyrAspAsn 208
 DB 437 ACATTCGTATGAAGTAAATGATTTATTAACAAACTAATGATGATGATGATGATGATGAT 496
 QY 209 AlaPheGlyValGluArgAlaGlnSerValThrGlyAsnIleLeuValCysSerGly 228
 DB 497 GGTTCCTGCTGAGCGTGCAGCACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 556
 QY 229 ProLeuSerValThrArgGluValValValProAsnIleAspArgTyrIleAsnGln 248
 DB 557 CGGTAAAGTGTATGCTAGAGAAAGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 616
 QY 249 ThrPheLeuGlyIleProValSerIleGlyAspAspAlaGlnThrAsnTyrAlaThr 268
 DB 617 ATGTCCTGCTGAGAGGTGAGTGTGGAGATGATGATGATGATGATGATGATGATGATGAT 676
 QY 269 AspLeuGlyLysThrValThrGlnSerThrAlaLysCysIleThrAspValProAspLys 288
 DB 677 TTGAAGGGAACACAGTTATCAATCCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
 QY 289 MetSerThrThrLeuLysGlnGlnAsnArgTyrAsnLysSerPhePheArgGluSerIle 308
 DB 737 TTAACAACATTTCTTAACAGCAACTACGTTGCAACAGTCAATTTTATGAGAAAGTTTA 796
 QY 309 IleSerValLysLysIleMetAsnPropheValAlaLeuThrThrIleLeuGluVal 328
 DB 797 ATTTCACTGGCATTTGGTATGAAAAACCAATGTTCTGTTGGACATTTTTCGAAATA 856
 QY 329 SerMetPheMetMetLeuValTyrSerValValAspPhePheValGlyAsnValArgGlu 348
 DB 857 TCGTTATGGATTTATTTGGGCTTTCCTTCAAGTAAATATTCATCAAGGCAAGTCAAT 916
 QY 349 PheAspTrpLeuArgValLeuAlaPheLeuValIleIlePheIleValAlaLeuCysArg 368
 DB 917 GTAGGGTTAATTTGGCTGTTTATTTATTTGGGTTATATTTATTTATTTATTTATTTAT 976
 QY 369 AsnIleHisTyrMetLeuLysHisProLeuSerPheLeuLeuSerProPheTyrGlyVal 388
 DB 977 AATGATTTATCTATTAACATCCCTTACTTCTTACTTCTTACTTCTTACTTCTTACTTCT 1036

QY 389 LeuHisLeuPheValLeuGlnProLeuLysLeuTyrSerLeuPheThrIleArgAsnAla 408
 DB 1037 CTCATGATTAGCACTATTACCTATAGCGTTTATGCTTAACTAAGTAAATCTAAT 1096
 QY 409 AspTyrGlyThrArg 413
 DB 1097 GGTGGGGAACAGT 1111
 RESULT 14
 US-09-939-964-1/c
 ; Sequence 1, Application US/09939964
 ; Publication No. US20030054522A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosenthal, Andre
 ; APPLICANT: Freiberg, Christoph
 ; APPLICANT: Perret, Xavier Philippe
 ; APPLICANT: Broughton, William John
 ; TITLE OF INVENTION: Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic
 ; TITLE OF INVENTION: Plasmid
 ; FILE REFERENCE: CARP0068
 ; CURRENT APPLICATION NUMBER: US/09/939,964
 ; CURRENT FILING DATE: 2001-08-27
 ; PRIOR APPLICATION NUMBER: 09/214,808
 ; PRIOR FILING DATE: 1999-06-22
 ; NUMBER OF SEQ ID NOS: 1
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 536165
 ; TYPE: DNA
 ; ORGANISM: Rhizobium
 ; US-09-939-964-1
 Alignment Scores:
 Pred. No.: 6,37e-39 Length: 536165
 Score: 463.00 Matches: 131
 Percent Similarity: 48.75% Conservative: 64
 Best Local Similarity: 32.75% Mismatches: 185
 Query Match: 21.59% Indels: 20
 DB: 11 Gaps: 10
 US-09-469-200D-2 (1-417) x US-09-939-964-1 (1-536165)
 QY 27 LeuPheGlyAlaLysGly-----SerLeuSerIleTyrGlyPheLeuLeuIleAlaTyr 44
 DB 157330 CTGCTTGGCAGCAGCGCGCTAGCCATCTCTGTATGACAGCAGCTCTCGACGGCTTAC 157271
 QY 45 -----LeuLeuValLysMetSerLeuSerPhePheTyrLysProPhe 58
 DB 157270 AAAGCGATGCAAGCTATATACGCTTGGCAACACACACAGCTGCTCAAGCGCGGTG 157211
 QY 59 LysGlyArgAlaGlyGlnTyrLysValAlaAlaIleIleProSerTyrAsnGluAspAla 78
 DB 157210 ACCGGCTCGGTGACCGCGAGCGGTGATGATGATGATGATGATGATGATGATGATGAT 157151
 QY 79 GluSerLeuLeuGluThrLeuLysSerValGlnGlnGlnThrTyrPro---LeuAlaGlu 97
 DB 157150 CGCGCGCTCTCGCGCTGCTGCTTCCATTCCTGCAAGCAGGAGTACGCTGAGAGTGGCG 157091
 QY 98 IleTyrValValAspAspGlySerAlaAspGluThrGlyIleLysArgIleGluAspThr 117
 DB 157090 GTCTAGCTGGTGTGACGAGCTTCTGCAATCGCAACGCCATCATACCTGTACACGATCAT 157031
 QY 118 ValArgThrGlyAspLeuSerSerAsnValIleValHisArgSerGluLysAsnGln 137
 DB 157030 TAT-----GGTGGCGAGCGGTGCTTCTTATCTG-----ATGCCAAAGAACGTC 156983
 QY 138 GlyLysArgHisAlaGlnAlaIlePalaPheGluArgSerAspAlaAspValPheLeuThr 157
 DB 156982 GAAAGCGCAAGCAGATGTCGATACGGGAATCATCGGAGATTTGGTGTCTCAAC 156923
 QY 158 ValAspSerAspThrTyrIleTyrProAspAlaLeuGluLeuLeuLysThrPheAsn 177

```
Db 156922 GTTGACTCGGACGACGACCATTTGGCGCGGAGCTAGTCAAGAAATTCGCCGTGAAGATGTAC 156863
QY 178 AspProThrValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeu 197
Db 156862 AGTCCCGGGTGGCGGCGGAGTGTGAGTTCAGCGGCCACACACCGCAGCACATGG 156803
QY 198 LeuThrArgLeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGln 217
Db 156802 CTGACGCGGTGATGCACATGAGTACTGGCTGCCTGCAACGAGGAAACGCGCACACAG 156743
QY 218 SerValThrGlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgGluVal 237
Db 156742 GTCGCTTTGGACCGCTATGTGTGTGGCGCGCTGTGCGCATGTACCGGGGCGCGCA 156683
QY 238 ValValProAsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIle 257
Db 156682 CTCCTATTGTCTGCTCATATAATACGACGACCACTGTTCGAGCGGCAACGACACTTC 156623
QY 258 GlyAspAspArgCysLeuThrAsnTyrAlaThrAspLeuGly---LysThrValTyrGln 276
Db 156622 GGGGAAGACCGCCACCTCAATCTCATCTGATGAGGCTTTCGAACCGAGTACGTT 156563
QY 277 SerThrAlaLysCysIleThrAspValProAspLysMetSerThrIleIleLysGlnGln 296
Db 156562 CCGGAAGCCATCGCGGCGGAGCTGTCTCCAAACTCGATGGGCGCTATCTGCGCCACAA 156503
QY 297 AsnArgTrpAsnLysSerPheArgGluSerIleIleSerValLysIleMetAsn 316
Db 156502 CTGCGCTGGGACGACGACGTTTCGGACACATGCTCGGCTCCGCTACTCGCGGCG 156443
QY 317 ---AsnProPheValAlaLeuThrPheIleLeuGluValSerMetPheMetMetLeuVal 335
Db 156442 CTTGATCGCTACTTCTAGCTGACGTGACGTGACGACAACTTGTGCTCGCTCTAGCC 156383
QY 336 TyrSerValValAspPhePhe-----ValGlyAsnValArgGluPheAspTrpLeuArg 353
Db 156382 CTTCTGGTCTGAGGGGCTAGCACGCTGCTGCGGCGCACAGTGCCTGGTCGAGC 156323
QY 354 ValLeuAlaPheLeuValIlePheIleValAlaLeuCysArgAsnIleHisTyrMet 373
Db 156322 ATCCCTG-----ATGATTGCATCTATGACAACTGTCCTGCTGCGTGC---GGCGTGGCGGCTT 156272
QY 374 LeuLysHisProLeuSerPheLeuLeuSerPropheTyrGlyValLeuHisLeuPheVal 393
Db 156271 CGAGGCGGAGCTGCGATCTCTTGGGTTTTCGCTGACACACCTCTCCACAGTCGCTCTC 156212
QY 394 LeuGlnProLeuLysLeuTyrSerLeuPheThrIleArgAsnAlaAspTrpGlyThrArg 413
Db 156211 CTGCTCCCTCAAGCATATGCGTGTGTGCACGTTGACGACAGGACTGCTGTGCGGT 156152
```

RESULT 15

```
US-09-902-939-3
; Sequence 3, Application US/0902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHavza
; APPLICANT: William Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/08020-USO
; CURRENT APPLICATION NUMBER: US/09/902-939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: mouse
US-09-902-939-3
```

```
Alignment Scores:
Pred. No.: 1,78e-40 Length: 1665
Score: 440.00 Matches: 134
Percent Similarity: 44.44% Conservative: 86
```

```
Best Local Similarity: 27.07% Mismatches: 171
Query Match: 20.51% Indels: 104
DB: 11 Gaps: 18

US-09-469-200D-2 (1-417) x US-09-902-939-3 (1-1665)

QY 4 LeuLysAsnLeuIleThrValValAlaPheSerIlePhe-----Trp 17
Db 13 CTGACTACAGCCCGCGTGTGGTGGCCACCACTGTGTTCCTCCCTGGTAGTCTGGGAGGC 72
QY 18 ValLeuLeuIleTyrValAsnValTyrLeuPhe-----GlyAlaLysGlySerLeuSer 35
Db 73 ATCCTGGCGGCTATGTGTGACAGGCTACCACTTATCCACACAGAAAGACACTACTCTGTC 132
QY 36 -----IleTyrGlyPheLeuLeuIleAlaTyrLeuLeuValLysMetSerLeuSerPhe 53
Db 133 TTGGGCTCTACGGTGGCCCTCTGCGGTCTACATCTGCTCATCCAGAGCCTGTTTGGCTTC 192
QY 54 PheTyrLysProPheLysGlyArgAlaGlyGln-----64
Db 193 CTGGAGCACCGTGCATGCGCAGGCGGCGCCCTCAAGTGCACCTGCTGCCACAGG 252
QY 65 ---TyrLysValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSerLeuLeuGlu 83
Db 253 TCAGCTTCAGTGGCCTCTGCTGCTTACCAAGAGGAGCCCGCAATACCTGCGCAAG 312
QY 84 ThrLeuLysSerValGlnGlnThrTyrProLeuAlaGlnIleTyrValValAspAsp 103
Db 313 TGCCTTCGCTCAGTCACGCGCATTCGCTTCCAAACCTCAAGGTGCTGTGTAGTGGAT 372
QY 104 GlySerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgAspThr-----121
Db 373 GGCAATCGCCAGGAA-----GATACCTACATGTTGGACATCTTCCAT 414
QY 122 -----GlyAspLeuSerSerAsnValIleValHisArgSer-----Glu 134
Db 415 GAGGTGCTGGTGGCGCTGAGACAGCTGCTTCTTGTGTGGCTAGCAATTTCCATGAG 474
QY 135 LysAsnGlnGlyLysArgHisAla-----142
Db 475 CGGGTGAAGAGAGACAGACAGGCGCCACCTGCAGGAGCATGGAGCGTGTCCGAGCTGTG 534
QY 143 -----GlnAlaTrpAlaPheLysArg-----149
Db 535 GTGTGGGCGCAGCACCTTCTCATCATCATGAGTGGGGGGGCGGCAAGCGTAGGCTCAG 594
QY 150 -----SerAspAlaAspValPheLeuThrValAspSerAsp 161
Db 595 TACACTGCCTTCAAGGCCCTTGGCACTCAGTGCAGTACATCCAGGTGTGTGACTCTGAC 654
QY 162 ThrTyrIleTyrProAspAlaLeuGluLeuLeuLysThrPhe---AsnAspProThr 180
Db 655 ACTGTGTCGACCCCGCTGACCATTTGAGATGCTTCGAGTCTTGGAGAGAGATCCCAAA 714
QY 181 ValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArg 200
Db 715 GTAGGAGTGTGGAGGAGATGTCCAAATCTCCACCAAGTATGATTCATGAGATCTCTTC 774
QY 201 LeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThr 220
Db 775 CTGAGCAGTGTGAGGTACTGGATGGCTTTCAACGTGGGCGGCGGCTGCGCACTCTACTTT 834
QY 221 GlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgGluValValPro 240
Db 835 GGCTGTGCAATGATATTAGTGGCGCTTTGGGCATGTACCGCAACAGGCTCTCTTCAGCAG 894
QY 241 AsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAsp 260
Db 895 TTCCTGGAGGATGGTACCATCAGAGTTCCTTAGGAGCAAGTGCAGCTTTGGGGATGAT 954
QY 261 ArgCysLeuThrAsnTyrAlaThrAspLeuGly---LysThrValTyrGlnSerThrAla 279
Db 955 CGGCACCTTACCACCGGAGTCTGTGCTTGGCTACCGGACTAAGTATATACAGACGCTCT 1014
```

```
QY 280 LysCysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrp 299
Db 1015 AAGTCCCTCACAGACCCCACTAGGTACCTTCGATGGCTCAATCAGCAACCCGCTGG 1074
QY 300 AsnLysSerPheArgGluSerIle----- 308
Db 1075 AGCAAGTCTTACTTCGGGAATGGCTCTACAATTCTGTGTGGTTCCATAGCACCACCTC 1134
QY 309 ---IleSerValLysIleMetAsn-----ProPheValAlaLeuTrpThrIle 325
Db 1135 TGGATGACCTATGAATCAGTGGTCACAGGTTCTTCCCATTTCTCCCTCATGTGTACAGTC 1194
QY 326 LeuGluValSerMetPheMetLeuValTyrSerValValAspPheValGlyAsn 345
Db 1195 ATACAACTT-----TTCACCGTGGCCGC 1218
QY 346 ValArgGluPheAspTrpLeuArgValLeuAlaPheLeuValIleIlePheIleValAla 365
Db 1219 ATC-----TGG---AACATTCTCTCTCTGCTAACAGTGCAGTGGTGGGC 1263
QY 366 LeuCysArgAsnIleHis---TyrMetLeuLysHisProLeuSerPheLeuLeuSerPro 384
Db 1264 ATATCAGGCTACCTATGCTGCTCTTCGAGCAATGCAGAGATGATCTTCATGTCC 1323
QY 385 PheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyrSerLeuPheThr 404
Db 1324 CTCTACTCCCTCTCTATATGTCCAGCCTCTTCCAGCCCAAGATCTTGTCTATTGCTACC 1383
QY 405 IleArgAsnAlaAspTrpGlyThr-----ArgLysLysLeuLeu 417
Db 1384 ATCAACAAGTCTGGTGGGCACCTCTGGCAGGAAACCAATTGTC 1428
```

Search completed: October 2, 2003, 16:20:50
Job time : 534 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 2, 2003, 12:39:19 ; Search time 2815 Seconds
(without alignments)
3600.343 Million cell updates/sec

Title: US-09-469-200D-2
Perfect score: 2145
Sequence: 1 MTKLNLTVAEISFWLL.....KLYSLTIRNADGTRKLL 417

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_p2n.model -DEV=xlh
-Q/cgn2.1/USPTO.spool/US09469200/runat_02102003_090348_22788/app_query.fasta.1.583
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -IOOCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MAFEX=biosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09469200.cgn.1.2810 @runat_02102003_090348_22788 -NCPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estm:*
5: em_estov:*
6: em_estpi:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gssi:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	440	20.5	4151	11	AK028582	Mus muscu
2	439	20.5	2891	11	AK079729	Mus muscu
3	320	14.9	809	14	CA324064	UI-M-FY0-
4	313.5	14.6	980	29	BZ606767	WHAD10TF
5	297.5	13.9	711	13	B0278152	B0278152
6	291	13.6	1084	12	BMS44718	AGENCOURT
7	286	13.3	915	13	EX371636	BM544718
8	284	13.2	639	9	AL775141	BM544718
9	280	13.1	551	10	BG707429	AL775141
10	280	13.1	551	10	BG707429	BG707429
11	273.5	12.8	656	12	B041963	BG707429
12	273	12.7	536	12	B062498	B041963
13	267.5	12.5	954	13	B0910636	B062498
14	263	12.3	866	14	CA988122	B0910636
15	262.5	12.2	876	13	B0907404	CA988122
16	261.5	12.2	782	12	B1753116	B0907404
17	260	12.1	857	13	BX327795	B1753116
18	259.5	12.1	948	13	BX327795	BX327795
19	245	11.4	623	9	AL966439	BX327795
20	242.5	11.3	672	12	B073147	AL966439
21	236	11.0	546	10	BF193947	B073147
22	224.5	10.5	699	9	AL647479	BF193947
23	222	10.3	643	28	A2282684	AL647479
24	215.5	10.0	883	14	CA793387	A2282684
25	199.5	9.3	704	12	B066629	CA793387
26	189	8.8	481	2	BSM089391	B066629
27	187.5	8.7	517	10	BG348186	BSM089391
28	186.5	8.7	542	10	BG364756	BG348186
29	185	8.6	1069	29	CNS0682W	BG364756
30	184.5	8.6	991	13	BQ725516	CNS0682W
31	182	8.5	923	13	B0909569	BQ725516
32	181.5	8.5	885	14	CA987219	B0909569
33	178	8.3	1135	29	CNS06522	CA987219
34	177.5	8.3	376	9	AU244838	CNS06522
35	176	8.2	895	13	B0911059	AU244838
36	173.5	8.1	548	10	BG347935	B0911059
37	170	7.9	862	29	CNS03W09	BG347935
38	169	7.9	859	14	CA986884	CNS03W09
39	164	7.6	485	13	BQ828135	CA986884
40	162	7.6	482	12	BG892747	BQ828135
41	155.5	7.2	869	14	CD052715	BG892747
42	155.5	7.2	941	29	CNS069FT	CD052715
43	154.5	7.2	533	9	AL924720	CNS069FT
44	154	7.2	750	28	BH367878	AL924720
45	153.5	7.2	831	14	CA986525	BH367878

ALIGNMENTS

RESULT 1
AK028582
LOCUS
DEFINITION
Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:4732404L04 product:similar to D642111 [Brachydanio rerio], full insert sequence.
ACCESSION
AK028582
VERSION
AK028582.1
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

- 1 Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
- 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
- 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakauchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system-384 format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
- 4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleschmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Hashio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, I., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weltz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409 (6821), 685-690 (2001)
 MEDLINE 21085660
 PUBMED 11217851
- 5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 MEDLINE 12141511
- 6 (bases 1 to 4151)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hayashizaki, Y., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akshira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in RIKEN contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/
 URL: http://fantom.gsc.riken.go.jp/.

COMMENT

FEATURES

SOURCE

1. 4151
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM_DB:4732404L04"
 /db_xref="taxon:10090"
 /clone="4732404L04"
 /tissue_type="Skin"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="10 days neonate"
 118. .1782

CDS

/note="unnamed protein product; putative similar to D642111 [Brachydanio rerio] (SPTR|Q9DG40, evidence: FASTA, 75.1%ID, 99.8%length, match=1659);"
 /codon_start=1
 /protein_id="BAC26017.1"
 /db_xref="GI:26324526"
 /translation="MPVLTALRVVGTSLFVLLVGLGILAAVTVGVQFIHTEKHVLS FGLVGLHLLIQSLFAFLHRRMERACRPIKLHCSQRPSPRVALCIIAYQEDPEVL RCLSEARAFAPNLYVMVVDGROEDYMLDIEHYLGGTQAGFFVWRSNPHEAG EETVASELQSGMERVAVYVNAFTSCIMKNGKREVMITAFALGNSVDYIQVCSGD TYVDPACTEMLRVLEEDPQVGGGVQILNKYDSWISFLSVRYVMAFNVACROS YFGVCQISGLPMVNSLQQLFEDYHQFLGSKSGDRLHLYNVLISGVTXY TARSCTETPRYLNQQLRWKSYFREWLYNSLWFKHLMWTYVSVTGFYF FLIAVQLQYFGRINWILLFLLVQLVQIKATYACFLGNAMFLMSLYSLVMS LPKAFIATATINKSGMTSGRKTIVNFGILPVSIVWVLLGLLAYTAYQDLFSE LLEAFVLSGALLYGCYVWALLMLYLIIARRCGKPEQYSIAFVY"
 4131. 4136
 /note="putative"
 4151
 /note="putative"

polyA_signal

polyA_site

BASE COUNT 983 a 1013 c 1055 g 1100 t

ORIGIN

Alignment Scores:
 Pred. No.: 1.39e-36 Length: 4151
 Score: 440.00 Matches: 134
 Percent Similarity: 44.44% Conservative: 86
 Best Local Similarity: 27.07% Mismatches: 171
 Query Match: 20.51% Indels: 104
 DB: 11 Gaps: 18

US-09-469-200D-2 (1-417) x AK028582 (1-4151)

QY 4 LeuLysAsnLeuLeuThrValValaLabeSerIlePhe-----Ttp 17
 130 CTGACTACAGCCGCGCTGGTGGGCGCACAGTCTGTTGCCCTGTAGTCGTGGAGGC 189
 18 ValLeuLeuLeuTyrValAsnValTyrLeuPhe-----GlyAlaLysGlySerLeuSer 35
 190 ATCCTGGCGGCGCTATGTGACAGCTACCGATTATTCACACAGAAAGCACTACCTGTCC 249
 36 -----IleTyrGlyPheLeuLeuLeuAlaTyrLeuLeuValLysMetSerLeuSerPhe 53
 250 TTGTGGCCTCTACGTGGCGCATCTCGGTGCTACATCTGCTCATCCAGAGCGCTTTGCTTC 309
 54 PheTyrLysProPheLysGlyAigAlaGlyGln----- 64
 310 CTGGACACCCGTCGCAATGCCAGGCGGCGGCCCTCAAGCTGCATGCTCCAGAGG 369

QY 65 ---TyrLysValAlaAlaIleIleProSerTyrAsnGluAspAlaGluSerLeuLeuGlu 83
 Db 370 CGCGCTTCAGTGGCATCTGCTGCTCCCTACCAAGAGGACCCCGAATACCTGGCGAAG 429
 QY 84 ThrLeuLysSerValGlnGlnInThrTyrProLeuAlaGluIleTyrValValAspAsp 103
 Db 430 TGCCTGCTCAGCTCAGCGCATTCCTCCAAACCTCAAGGTGGTCATGTTAGTGGAT 489
 QY 104 GlySerAlaAspGluThrGlyIleLysArgIleGluAspTyrValArgAspThr 121
 Db 490 GGCATGCCAGGAA-----GATACCTACATGTTGGACATCTTCCAT 531
 QY 122 -----GlyAspLeuSerSerAsnValIleValHisArgSer-----Glu 134
 Db 532 GAGTGTGGTGGCGACTGAGCAGCTGCTCTTGTGTGGCTACCAATTTCCATGAG 591
 QY 135 LysAsnGlnGlyLysArgHisAla----- 142
 Db 592 GCGGTGAAGGAGACAGAGCCGCTGCAGGAAGCATGGAGCGTGTGGAGCTGTG 651
 QY 143 -----GlnAlaTyrPalapheGluArg----- 149
 Db 652 GTGTGGCGCAGCACCTTCATGATCATCAGAAAGTGGGGGCGCAGGCTGAGTATG 711
 QY 150 -----SerAspAlaAspValPheLeuThrValAspSerAsp 161
 Db 712 TACACTGCCTTCAAGCCCTTGGCACTCAGTGGACTACATCCAGGTGTGCTCTGAC 771
 QY 162 ThrTyrIleTyrProAspAlaLeuGluLeuLysThrPhe---AsnAspProThr 180
 Db 772 ACTGTGTGCAGCCCTGCACCATTCAGATTCGATGCTTCGAGTCTTGGAAAGATCCCAA 831
 QY 181 ValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArg 200
 Db 832 GTAGAGGTGTGGAGAGATGCTCAATCTCCACAGATGATGATCATGGATCCCTTC 891
 QY 201 LeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThr 220
 Db 892 CTGAGCAGTGTGAGTACTGGATGGCTTTCACGTGGAGCGGCTGCCAGTCTACTTT 951
 QY 221 GlyAsnIleLeuValCysSerGlyProLeuSerValTyrArgArgGluValValPro 240
 Db 952 GCGCTGTGCAATGATTATAGTGGCCCTTGGCATGTACCCACAGCCTCTCCAGCAG 1011
 QY 241 AsnIleAspArgTyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAsp 260
 Db 1012 TTCTGGAGGATGGTACCATCAGAGTCTCTAGGACGCAAGTGCACCTTTGGGATGAT 1071
 QY 261 ArgCysLeuThrAsnTyrAlaThrAspLeuGly---LysThrValTyrGlnSerThrAla 279
 Db 1072 CGGCACCTTACCACCGAGTCTGAGTCTTGGCTACCGAGTAAAGTATACAGCGCTCT 1131
 QY 280 LysCysIleThrAspValProAspLysMetSerThrTyrLeuLysGlnAsnArgTyr 299
 Db 1132 AGTGCTTCACAGAGACCCCACTAGTACCTTCGATGGCTCAATCAGCAACCCGCTGG 1191
 QY 300 AsnLysSerPhePheArgGluSerIle----- 308
 Db 1192 ACCAAGTCTTACTTCGGGAATGGCTCTACAAATCTCTGTGTCCATAGCACCACTC 1251
 QY 309 ---IleSerValLysLysIleMetAsnAsp-----ProPheValAlaLeuThrPThrIle 325
 Db 1252 TGGATGACCTATGAATCAGTGTGTCACAGGTTCTTCCATCTCTCCATCTACAGTC 1311
 QY 326 LeuGluValSerMetPheMetMetLeuValTyrSerValValAspPhePheValGlyAsn 345
 Db 1312 ATACAACTT-----TTTACCGTGGCGCG 1335
 QY 346 ValArgGluPheAspThrLeuArgValLeuAlaPheLeuValIleIlePheIleValAla 365
 Db 1336 ATC-----TGG---NACATTCTCTCTCTCAACAGTGCAGTGTGGTGGC 1380

QY 366 LeuCysArgAsnIleHis---TyrMetLeuLysHisProLeuSerPheLeuLeuSerPro 384
 Db 1391 ATTATCAAGGCTACTAGCTCTCTCTTCGGAAGTGCAGAGATGATTCATGTCTC 1440
 QY 385 PheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTyrSerLeuPheThr 404
 Db 1441 CTCTACTCCCTCTCTATATATGTCAGCCTTTCGCCAGCAAGATCTTGTCTATGTCTAC 1500
 QY 405 IleArgAsnAlaAspTyrPThrGlyThr-----ArgLysLysLeuLeu 417
 Db 1501 ATCAACAAGTCTGGCTGGGCACTTCTGCGAGGAACCAATTTGTC 1545
 RESULT 2
 LOCUS AK079729 2891 bp mRNA linear HTC 05-DEC-2002
 DEFINITION Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched library, clone:A430041J01 product:hyaluronan synthase 2, full insert sequence.
 ACCESSION AK079729
 VERSION 1 GI:26348186
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20493374
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 REFERENCE 4
 AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Washima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J. H., Kohsaki, S. and Hayashizaki, Y.
 TITLE Functional annotation of a full-length mouse cDNA collection

Db 1500 CAACAGGTTGTGAGTCTGGCTATGCAACTAAATACAGCGCTGGTCCAAAGTGCCTTAC 1559
 QY 283 rAspValProAspLysMetSerThrTyrLeuLysGlnAsnArgTrpAsnLysSerPh 303
 Db 1560 TGAACCTCCATAGATATCTCAGATGGCTGAACACAGACACCGTTGGAGCAAGTCCCTA 1619
 QY 303 ePheArgGluSerIleLeuSerValLysLysLysLysLysLysLysLysLysLysLys 323
 Db 1620 CTTCCGAGAGTGGCTGTACAAATGCCATGGTGTTCACAGCAT-----CACTTGTG 1670
 QY 323 pThrIleuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 338
 Db 1671 GATGACCTATGAAGCTGTATCATCAGGATCTTTTCCTTTCTTCATTCGCCACAGTAT 1730
 QY 338 lValAspPhePheValGlyAsnValArgGluPheAspTrpLeuArgValLeuAlaPheLe 358
 Db 1731 CCAGCTCTTCTACAGGGGTAAATC-----TGG-----AACATCTCTCTCTCTCT 1775
 QY 358 uValIleIlePheLeuValAlaLeuCysArg---AsnIleHisTyrMetLeuLysHisPr 377
 Db 1776 GTTAACCTGCCAGTAGTGGTCTCATCAAGTCATCTTTTGGCAGCTGCCTTAGAGGAAA 1835
 QY 377 cLeuSerPheLeuLeuSerProPheTyrGlyValLeuHisLeuPheValLeuGlnProLe 397
 Db 1836 TATCGTCATGATTCATCT 1895
 QY 397 uLysLeuTyrSerLeuPheThrIleArgAsnAlaAspTrpGlyThr-----ArgLysLy 415
 Db 1896 CRAAGTGTTCATTCATTCACCAATCAACAAAGCTGGGTGGGCACATCTCGAAGGAGAC 1955
 QY 415 sLeuLeu 417
 Db 1956 CATTTGTT 1962
 RESULT 3
 CA324064
 LOCUS
 DEFINITION UI-M-FY0-cco-g-16-0-01.r1 NIH_BMAP_FY0 Mus musculus CDNA clone
 IMAGE: 6822281 5', mRNA sequence.
 CA324064
 ACCESSION CA324064.1 GI:24542162
 VERSION
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 809)
 NIH-MGC <http://mgi.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP).

FEATURES
source

Seq primer: pYX-5.
 Location/Qualifiers
 1..809
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE: 6822281"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5, 14.5, 16.5, 17.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_FY0"

/note="Organ: Brain; Vector: pYX-Asc; Site: 1; Ecor I;
 Site: 2; Not 1; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with Ecor I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is ACGGACAG. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemlin Chin, Ph.D.,
 program coordinator."

BASE COUNT 175 a 218 c 193 g 221 t 2 others
 ORIGIN

Alignment Scores:

Pred. No.: 1,94e-24 Length: 809
 Score: 320.00 Matches: 76
 Percent Similarity: 50.57% Conservative: 57
 Best Local Similarity: 28.90% Mismatches: 96
 Query Match: 14.92% Indels: 34
 DB: 14 Gaps: 8

US-09-469-200D-2 (1-417) x CA324064 (1-809)

QY 166 ProAspAlaLeuGluLeuLeuLysThrPhe---AsnAspProThrValPheAlaAla 184
 Db 12 CCAGCTGCACCATGAGATGCTTCGAGCTTTGGAGAAGATCCCAAGTAGGAGTGT 71
 QY 185 ThrGlyHisLeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArgLeuThrAspIle 204
 Db 72 GGAGGAGATGCTCAATCTCTCAACAGTATGATTCATGGATCTCTCTCTGAGCAGTGTG 131
 QY 205 ArgTyrAspAlaPheGlyValGluArgAlaGlnSerValThrGlyAsnIleLeu 224
 Db 132 AGGTACTGGATGGCTTCAACGTGGAGCGGCGCTCCAGCTCTACTTTGGCTGTGCAC 191
 QY 225 ValCysSerGlyProLeuSerValTyrArgArgGluValValValProAsnIleAspArg 244
 Db 192 TGTATTAGTGGCGCTTTGGCATGTACCGCAACAGCTCTCTTCAGCAGTCTCTGGAGAT 251
 QY 245 TyrIleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspArgCysLeuThr 264
 Db 252 TGGTACCATCAGAAGTCTCTAGGAGCAGATGCAGCTTTGGGATGATCGGACCTTACC 311
 QY 265 AsnTyrAlaThrAspLeuGly---LysThrValTyrGlnSerThrAlaLysCysIleThr 283
 Db 312 AACCGAGTCTCTGAGTCTTGGTACCGGACTAAGTATATACAGCAGCGCTTAAGTGCCTACA 371
 QY 284 AspValProAspLysMetSerThrTyrLeuLysGlnAsnArgTrpAsnLysSerPhe 303
 Db 372 GAGACCCCACTAGTACTCTTCGATGGCTCATCATCAGAAACCCGCTGGAGCAGTCTTAC 431
 QY 304 PheArgGluSerIle-----IleSerVal 311
 Db 432 TTTCCGGAATGGCTCTCAACAATTCCTGTGTGTTCCATAAGCACCACTCTGGATGACCTAT 491
 QY 312 LysLysIleMetAsnAsn-----ProPheValAlaLeuThrPheLeuGluValSer 329
 Db 492 GAATCAGTGGTCACAGGTTTCTTCCCATTTCTTCCTCATTTGCTACATCATCAACTT--- 548
 QY 330 MetPheMetMetLeuValTyrSerValValAspPheValGlyAsnValArgGluPhe 349
 Db 549 -----TTCACCGTGGCGCATC----- 566
 QY 350 AspTrpLeuArgValLeuAlaPheLeuValIleIlePheIleValAlaLeuCysArgAsn 369
 Db 567 ---TGG---AACATCT 620
 QY 370 IleHis---TyrMetLeuLysHisProLeuSerPheLeuLeuSerProPheTyrGlyVal 388

```

Db 621 ACCTATGCGCTTCTTCGAGCAATGCAGAGATGATCTTCATGTCNCCTCTACTCCCTT 680
    |||
Qy 389 LeuHisLeuPheValLeuGlnProLeuLysLeuThrSerLeuPheThrIleArgAsnAla 408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 681 CTCATATGTCAGCGCTTTCGAGCAATGATCTTCTATGCTTCTACCATCAACAAGTCT 740
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 409 AspTrpGly 411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 741 GGCTGGGA 749

RESULT 4
BZ606767
LOCUS WHADE10TF Human MCF7 breast cancer cell line library (MCF7_1) Homo
DEFINITION sapiens genomic clone MCF7_1-20A19, genomic survey sequence.
ACCESSION BZ606767
VERSION BZ606767.1 GI:31515328
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 980)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 980)
Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
Gray,J.W. and Collins,C.
ESP: a sequence-based approach to the structural genomics of tumors
UNPUBLISHED (2002)
CONTACT: Volik SV
Colin Collins' lab
UCSF Comprehensive Cancer Center
UCSF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
Fax: 415 502 5665
Email: svolik@cc.ucsf.edu
This clone is available from Amplicon Express
http://www.genomex.com
Class: BAC ends.
FEATURES
    source
        1..980
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /clone="MCF7_1-20A19"
            /sex="female"
            /clone_lib="Human MCF7 breast cancer cell line library
            (MCF7_1)"
            /note="Vector: pCRAC1; Site_1: HindIII; This library was
            constructed from MCF7 breast cancer cell line by Amplicon
            Express (http://www.genomex.com) using their standard
            procedure."
BASE COUNT 257 a 198 c 215 g 310 t
ORIGIN
Alignment Scores:
Pred. No.: 1.33e-23 Length: 980
Score: 313.50 Matches: 76
Percent Similarity: 52.72% Conservative: 50
Best Local Similarity: 31.80% Mismatches: 96
Query Match: 14.62% Indels: 17
DB: 29 Gaps: 7

US-09-469-200D-2 (1-417) x BZ606767 (1-980)
Qy 188 LeuAsnValArgAsnArgGlnThrAsnLeuLeuThrArgLeuThrAspIleArgTrp 207
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 29 TTGCAGATTTTAAACAAGTACCATCTCTGGATCTATCTCCACAGAGTGTAGATATGG 88
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 208 AsnAlaPheGlyValGluArgAlaGlnSerValThrGlyAsnIleLeuValCysSer 227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 89 ATGGCTTTTAAATAGAAAGGCCGTGTCAGTCTTATTTGGTGTCTTTCAGTGCATTAGT 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Qy 228 GlyProLeuSerValTrpArgGluValValProAsnIleAspArgTrpIleAsn 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 149 GGACCTCTGGGATGATACAGAACTCTTGTTCATGAGTTTGGGAATGGTACAAAT 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 248 GlnThrPheLeuGlyIleProValSerIleGlyAspAspArgCysLeuThrAsnTrpAla 267
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 209 CAAGAATTTAGGCAACCAATGATGTTGGTGTAGCAGGAGCATCTCAGCAACCGGGTG 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 268 ThrAspLeuGly--LysThrValTrpGlnSerThrAlaLysCysIleThrAspValPro 286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 269 CTGAGCCTGGGCTATGCAACAAATACACAGCTCGATCTAAGTCCCTTACTGAAACACT 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 287 AspLysMetSerThrTrpLeuLysGlnAlaAsnArgTrpAsnLysSerPhePheArgGlu 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 329 ATAGATATCTCAGATGGGTAAACACGACAGCCGTTGGAGCAAGTCTCTACTTCCGAGA 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 307 SerIleLeuSerValLysLysIleMetAsnAsnProPheValAlaLeuThrPheIleLeu 326
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 389 TGGCTGTACATGCAATGTGGTTTCACAAACAT-----CACTTGTGGATGACCTAC 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 327 GluValSerMet-----PheMetMetLeuValTrpSerValValAspPhe 341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 440 GAAGCATATCTACTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 342 PheValGlyAsnValArgGluPheAspTrpLeuArgValLeuAlaPheLeuValIleIle 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 500 TACCGGGGTAAAT-----TGG---AACATTCCTCTCTCTCTCTCTCTCTCTCTCTCT 544
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 362 PheIleValAlaLeuCysArg---AsnIleHisTrpMetLeuLysHisProLeuSerPhe 380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 545 CAGCTAGTAGGTCTCATATAAATCATCTTTTGGCCAGCTGCTTAGAGGAAATATCGCTCATG 604
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 381 LeuLeuSerProPheTrpGlyValLeuHisLeuPheValLeuGlnProLeuLysLeuTrp 400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 605 GCTTCATGCTCTCTACTACTGCTGTTATACATGTCAGTTACTTCGCCGCAAGATGTTT 664
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 401 SerLeuPheThrIleArgAsnAlaAspTrpGlyThr-----ArgLysLysLeuLeu 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 665 GCATTTGCAACAATAACAAGCTGGTGGGCGACATCATCAGGAAGGAAACCATTTGT 721
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
BZ278152
LOCUS 603865094F1 CSEQCHN54 Gallus gallus cdna clone CHEST882124 5', mRNA
DEFINITION sequence.
ACCESSION BZ278152
VERSION BZ278152.1 GI:25727606
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
REFERENCE 1 (bases 1 to 711)
AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
)
PO Box 88, Manchester, M60 10D, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1..711
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
FEATURES
    source

```


Db 99 CAGGTCTGTCGACGACCAAGGTTGGACCCCGGAGCTGCTGGAGCTGCTGGCGGTA 158
 QY 176 PheAsn---AspProThrValPheAlaAlaThrGlyHisLeuAsnValArgAsnArgGln 194
 Db 159 CTGGACGAGGACCCCGGCTAGGCGCTGTTGGTGGGAGCGTGGGATCTTAACCCCTG 218
 QY 195 ThrAsnLeuLeuThrArgLeuThrAspLeuThrAspAsnAlaPheGlyValGluArg 214
 Db 219 GACTCTGGTGGTACCTTCTAGCAGCTCGGATAGCTGGTGGTGGTGGTGGTGG 278
 QY 215 AlaAlaGlnSerValThrGlyAsnLeuValCysSerGlyProLeuSerValThrArg 234
 Db 279 GCTTCTCAGAGCTACTTCCACTGTGTATCTCTGATCAGCGGTCTCTAGGCTATATAGG 338
 QY 235 ArgGluValValProAsnLeuAspArgTyrIleAsnGlnThrPheLeuGlyIlePro 254
 Db 339 ATATACCTCTGACGAGTTCTTCTAGGCTGGTATACACCAAGTTCTTGGGTATCCAC 398
 QY 255 ValSerIleGlyAspAspArgCysLeuThrAsnTyrAlaThrAspLeuGly---LysThr 273
 Db 399 TGTACTTTTGGGATGACCGGACCTCACCACCGCATGCTCAGCATGGGTATGCTACC 458
 QY 274 ValTyrGlnSerThrAlaLysCysIleThrAspValProAspLysMetSerThrTyrLeu 293
 Db 459 AAGTACACTTCCAGGCTCGGCTGCTACTCAGAGCGGCTGCTCTCTCTGCGGTGGCTG 518
 QY 294 LysGlnGlnAsnArgTyrAsnLysSerPhePheArgGluSerIle----- 308
 Db 519 ACCGACGACGACGCTGCTTCCAGTCTGCTACTTCCGTGCTGCTGCTGCTGCTGCTG 578
 QY 309 -----IleSerValLysLysIleMetAsnAsn-----ProPhe 319
 Db 579 TGGCACCAGGACCATCGCTGGATGACCTACGAGGCGGTGCTCCGCCCTGTTCGCCCTC 638
 QY 320 ValAlaLeuThrPheIleLeuValSerMetPheMetMetLeuValTyrSerValVal 339
 Db 639 TTCGTGGCGCCACTGTGCTGCTGCTG----- 665
 QY 340 AspPhePheValGlyAsnValArgGluPheAspTyrLeuArgValLeuAlaPheLeuVal 359
 Db 666 ---TTCATACGGCGCGCCCTTGGCGCTGCTGGGTGCTGCTGCTGCTGCTGCTGCTG 722
 QY 360 IleLeuPheLeuValAlaLeuCysArgAsnIleHisTyrMetLeuLysHisProLeuSer 379
 Db 723 GCACCTGGCGAAGCGGCTTCCGCGCTGCTGCGGCGCTGCTGCGGCGCTGCTGCTGCTG 782
 QY 380 PheLeuLeuSerProPheTyr-----GlyValLeuHisLeuPheValLeuGlnProLeu 397
 Db 783 TGTCTTACGGCGCCCTTCTACATGGGGGCGCTCTCTG-----CCTGCC 824
 QY 398 LysLeuTyrSerLeuPheThrIleArgAsnAlaAspTyrGlyThr-----ArgLysLys 415
 Db 825 AGTTCTTGGCGCTAGTCCCTGTAACCGGAGCGGTGGGCGCTTCCGCGCGCGGGA 884
 QY 416 Leu 416
 Db 885 CTG 887

RESULT 7

LOCUS BX371636
 DEFINITION BX371636 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CS0D1027J02 3-PRIME, mRNA sequence.
 ACCESSION BX371636
 VERSION BX371636.1
 KEYWORDS GI:30450029
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 915)
 AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
 TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 9502.r for
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0BA10192C12_CS01819_1&cluster=9502.r.
 Contact : Feng liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CS0BA10192C12_CS01819_1.
 FEATURES
 Location/Qualifiers
 1..915
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1027J02"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 181 a 302 c 268 g 155 t 9 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1.26e-20 Length: 915
 Score: 286.00 Matches: 73
 Percent Similarity: 50.70% Conservative: 36
 Best Local Similarity: 33.95% Mismatches: 86
 Query Match: 13.33% Indels: 20
 DB: 13 Gaps: 5
 US-09-469-200d-2 (1-417) x BX371636 (1-915)
 QY 119 ArgAspThrGlyAspLeuSerSerAsnValIleValHisArg----- 132
 Db 686 AAGATCTCTGGCGCGCTGGAGCGCTGGAGCGCTGAGGACTCGCAGGTGGTGGTGGTGG 627
 QY 133 SerGlyAsnGlnGlyLysArgHisAlaGlnAlaThrAlaPheGluArg-----Ser 150
 Db 626 GCACGCGCTGGCGCGCAACGCGGAGTGTATACACAGCTTCAGGCGCTCGGAGAT 567
 QY 151 AspAlaAspValPheLeuThrValAspSerAspThrTyrIleTyrProAspAlaLeuGlu 170
 Db 566 TCGGTGGACTACGTGCGAGGTCTGTGACTCGCACACAAGGTGGACCCCATGGCACTGCTG 507
 QY 171 GluLeuLeuLysThrPheAsn---AspProThrValPheAlaAlaThrGlyHisLeuAsn 189
 Db 506 GAGCTCTGGCGGTACTGGAG 447
 QY 190 ValArgAsnArgGlnThrAsnLeuLeuThrArgLeuThrAspIleArgTyrAspAsnAla 209
 Db 446 ATCTTAAACCTCTGGACTCTCTGGGTAGCTTCTTAAGCAGCGCTCGGATCTGGGTAGCC 387
 QY 210 PheGlyValGluArgAlaAlaGlnSerValThrGlyAsnIleLeuValCysSerGlyPro 229
 Db 386 TTCAATGTGGAGCGGCTTCTGAGAGCTTCTCCAGCTTCTCCAGCTGCTATCTGTCATC 327
 QY 230 LeuSerValTyrArgArgGluValValProAsnIleAspArgTyrIleAsnGlnThr 249
 Db 326 CTAGGCTTATAGGAATAACTCTTGTGACGAGGTTCTTGGAGGCTGTGTACACCAAGAG 267
 QY 250 PheLeuGlyIleProValSerIleGlyAspAspArgCysLeuThrAsnTyrAlaThrAsp 269
 Db 266 TTCTCTGGGTACCCACTCTACTTTTGGGATGACGCGGCACTCACCACCGCATGCTCAGC 207
 QY 270 LeuGly---LysThrValTyrGlnSerThrAlaLysCysIleThrAspValProAspLys 288
 Db 206 ATGGGTTATGTACCAAGTACACCTCCAGGTCCCGCTGCTCTACTCAGAGCGCCCTCGTCC 147

QY 289 MetSerThrTyrLeuLysGlnGlnAsnArgTirPAsnLysSerPhePheArgGluSerIle 308
 Db 146 TTCTCGGCTGGCTGAGCCAGACAGCTGGTCCAGTCGTACTTCCGTGAGTGGCTG 87
 QY 309 IleSerValLysLysIleMetAsnAsnProPheValAlaLeuTirP 323
 Db 86 -----TACACGCGCTCG 72

RESULT 8
 AL775141
 LOCUS AL775141 639 bp mRNA linear EST 25-JUN-2002
 DEFINITION AL775141 XGC-gastrula silurana tropicalis cDNA clone TGas077a10 5', mRNA sequence.

ACCESSION AL775141
 VERSION AL775141.1 GI:21560845
 KEYWORDS EST.
 SOURCE Silurana tropicalis (western clawed frog)
 ORGANISM Silurana tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Silurana.
 REFERENCE 1 (bases 1 to 639)
 AUTHORS Taylor, R., Ashurst, J.L., Croning, M.D.R., Zorn, A.M. and Rogers, J.
 TITLE Sanger Xenopus tropicalis EST project 2002
 JOURNAL Unpublished
 COMMENT Contact: Taylor R
 Sanger Centre
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 TROPICALIS_SEQUENCE ID: TGas077a10.plksp6
 Sequencing primer: plksp6
 This sequence is from a Xenopus Gene Collection (XGC) library constructed by Aaron M. Zorn.

FEATURES
 source
 1..639
 Location/Qualifiers
 /organism="Silurana tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="TGas077a10"
 /dev_stage="gastrula (stages 10.5-13 mixed)"
 /lab_host="Escherichia coli XL1-blue"
 /clone_lib="XGC-gastrula"
 /note="vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dt primed from 5ug of poly A+ RNA from stages 10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end."
 BASE COUNT 160 a 152 c 173 g 154 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,24e-20 Length: 639
 Score: 284.00 Matches: 68
 Percent Similarity: 52.06% Conservative: 33
 Best Local Similarity: 35.05% Mismatches: 79
 Query Match: 13.24% Indels: 14
 DB: 9 Gaps: 4

US-09-469-200D-2 (1-417) x AL775141 (1-639)

QY 134 GluLysAsnGlnGlyLysArgHisAlaGlnAlaTirPAlaPheGluArg-----SerAsp 151
 Db 13 CAAAATGGGGGAGAGAGAGAGTGTATGATACACAGCTTCAGGCACCTGGGGAGAC 72
 QY 152 AlaAspValPheLeuThrValAspSerAspThrTyrIleTyrProAspAlaLeuGluGlu 171
 Db 73 GTGGATATGTGTCAGGTATGCGACTCTGCACAGTGTGTCGCGCATCGCTGGTGGAG 132
 QY 172 LeuLeuLysThrPhe---AsnAspProThrValPheAlaAlaThrGlyHisLeuAsnVal 190
 Db 133 ATGCTCAAGTACTGGAGGAGAGATCATGCTTGGAGAGTGGGTGGAGATGTGCAGATT 192

QY 191 ArgAsnArgGlnThrAsnLeuLeuThrArgLeuThrAspIleArgTyrAspAsnAlaPhe 210
 Db 193 TTAACAAGTACGACTCGGTGGATTCTCTCTGAGTACGATACATGGATGGCGTTT 252
 QY 211 GlyValGluArgAlaAlaGlnSerValThrGlyAsnIleLeuValCysSerGlyProLeu 230
 Db 253 AACATGAGAGCGGTGCGAGTCTTACTCGGCTGTGTGCAATGCATCAGCGGCCCTTG 312
 QY 231 SerValTyrArgGluValValProAsnIleAspArgTyrIleAsnGlnThrPhe 250
 Db 313 GGGATGACCGGACTCTCTCCAGCAATCTCTGAGATGGTACACCAAGATTT 372
 QY 251 LeuGlyIleProValSerIleGlyAspArgCysLeuThrAsnTyrAlaThrAspLeu 270
 Db 373 TTGGGTTCCCGAGTGCAGTTTCCGGGATGATGCGCACCTCACCACCGAGTCTTGAGTCTG 432
 QY 271 Gly---LysThrValTyrGlnSerThrAlaLysCysIleThrAspValProAspLysMet 289
 Db 433 GGTACCGCAACCAATACACGCGCCAGTCCAAATGCTTACTGAAACCCCAACCGAGTAC 492
 QY 290 SerThrTyrLeuLysGlnGlnAsnArgTirPAsnLysSerPhePheArgGluSerIle 309
 Db 493 CTACGATGGCTCAACACGACAAACGCGGTGGAGCAAGTCTTACTCCGAGATGGCTG--- 549
 QY 310 SerValLysLysIleMetAsnAsnProPheValAlaLeuTirP 323
 Db 550 -----TACAATGCAATGGG 564

RESULT 9
 BG707429
 LOCUS BG707429 551 bp mRNA linear EST 07-MAY-2001
 DEFINITION 602672848f1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4795643 5', mRNA sequence.

ACCESSION BG707429
 VERSION BG707429.1 GI:13983769
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 551)
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toehiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM10678 row: p column: 12
 High quality sequence stop: 551.
 Location/Qualifiers
 1..551
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4795643"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_96"
 /note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcag)); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library

constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library.
 BASE COUNT 102 a 158 c 168 g 123 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,74e-20 Length: 551
 Score: 280.00 Matches: 69
 Percent Similarity: 51.58% Conservative: 29
 Best Local Similarity: 36.32% Mismatches: 78
 Query Match: 13.05% Indels: 14
 DB: 10 Gaps: 4

US-09-469-200d-2 (1-417) x BG707429 (1-551)

QY 138 GlyLysArgHisAlaGlnAlaThrPheGluArg-----SerAspAlaValPhe 155
 |||||
 Db 11 GCGAAGCCGAGGTGATACACAGCCCTCAGGGCGCTCGGAGATTCGGTGACCTACG 70
 QY 156 LeuThrValAspSerAspThrTyrIleThrProAspAlaLeuGluLeuLeuLysThr 175
 |||||
 Db 71 CAGGTCTGTGACTCGACACAAAGGTGGACCCATGCGACTCTGGAGCTCTGGGGTA 130
 QY 176 PheAsn---AspProThrValPheAlaThrGlyHisLeuAsnValArgGln 194
 |||||
 Db 131 CTGGACGAGGACCCCGGTAGGGCTGTGGTGGGACGTGCGGATCTTAACCTCTG 190
 QY 195 ThrAsnLeuLeuThrArgLeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArg 214
 |||||
 Db 191 GACTCTGGTGCAGCTCTCAGACGCTCGATACCTGGTACCTCAATGGGACGG 250
 QY 215 AlaAlaGlnSerValThrGlyAsnIleLeuValCysSerGlyProLeuSerValTyrArg 234
 |||||
 Db 251 GCTTGTGAGACTTCTCCATCTGTATCTCTGATCAGCGGTCTCTAGGCTATATAG 310
 QY 235 ArgGluValValProAsnIleAspArgTyrIleAsnGlnThrPheLeuGlyLeuPro 254
 |||||
 Db 311 ATATACCTTCCTGACGAGTTCCTTGGAGCCTGGTACACCAAGTTCCTGGGTACCC 370
 QY 255 ValSerIleGlyAspArgCysLeuThrAsnTyrAlaThrAspLeuGly---LysThr 273
 |||||
 Db 371 TGTACTTTGGGGTACCGGACCTCACCACCGATGCTCAGCAGCTGATGCTACC 430
 QY 274 ValTyrGlnSerThrAlaLysCysIleThrAspValProAspLysMetSerThrTyrLeu 293
 |||||
 Db 431 AGTACACCTCCAGGTCCTGCTACTCAGACGACCCCTGCTCTCTGCGGTGCTG 490
 QY 294 LysGlnGlnAsnArgTyrAsnLysSerPhePheArgGluSerIleIleSerValLysLys 313
 |||||
 Db 491 AGCCAGCAGACGCTGGTCCCAAGTCTGCTACTTCCGTGAGTGGCTG 535
 QY 314 IleMetAsnProPheValAlaLeuTyr 323
 |||||
 Db 536 -----TACACGCGCTCGT 550

RESULT 10

LOCUS BG707608 551 bp mRNA linear EST 07-MAY-2001
 DEFINITION 602670672P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4793704 5',
 mRNA sequence.

ACCESSION BG707608
 VERSION 1
 KEYWORDS BG707608.1 GI:13984125
 SOURCE EST.
 ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 551)
 AUTHORS NIH-MGC <http://mgi.nhl.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@email.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLNL0673 row: 0 column: 17
 High quality sequence stop: 551.
 Location/Qualifiers
 1..551
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:4793704"
 /tissue_type="hypothalamus"
 /lab_host="DH103"
 /clone_lib="NIH_MGC_96"
 /note="Organ: brain; Vector: pBluescriptR (modified
 pBluescript KS+); Site: 1: BamHI; Site: 2: SalI-XhoI (gtcgag
); Oligo-dt primed using primer 5'-rrrrrrrrrrrrrrrrrrrr-3',
 size-selected for average insert size 2.3 kb and
 normalized to 500. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."

FEATURES
 source

BASE COUNT 102 a 158 c 168 g 123 t
 ORIGIN

Alignment Scores:

Pred. No.: 2,74e-20 Length: 551
 Score: 280.00 Matches: 69
 Percent Similarity: 51.58% Conservative: 29
 Best Local Similarity: 36.32% Mismatches: 78
 Query Match: 13.05% Indels: 14
 DB: 10 Gaps: 4

US-09-469-200d-2 (1-417) x BG707608 (1-551)

QY 138 GlyLysArgHisAlaGlnAlaThrPheGluArg-----SerAspAlaValPhe 155
 |||||
 Db 11 GCGAAGCCGAGGTGATACACAGCCCTCAGGGCGCTCGGAGATTCGGTGACCTACG 70
 QY 156 LeuThrValAspSerAspThrTyrIleThrProAspAlaLeuGluLeuLysThr 175
 |||||
 Db 71 CAGGTCTGTGACTCGACACAAAGGTGGACCCATGCGACTCTGGAGCTCTGGGGTA 130
 QY 176 PheAsn---AspProThrValPheAlaThrGlyHisLeuAsnValArgGln 194
 |||||
 Db 131 CTGGACGAGGACCCCGGTAGGGCTGTGGTGGGACGTGCGGATCTTAACCTCTG 190
 QY 195 ThrAsnLeuLeuThrArgLeuThrAspIleArgTyrAspAsnAlaPheGlyValGluArg 214
 |||||
 Db 191 GACTCTGGTGCAGCTCTCAGACGCTCGATACCTGGTACCTCAATGGGACGG 250
 QY 215 AlaAlaGlnSerValThrGlyAsnIleLeuValCysSerGlyProLeuSerValTyrArg 234
 |||||
 Db 251 GCTTGTGAGACTTCTCCATCTGTATCTCTGATCAGCGGTCTCTAGGCTATATAG 310
 QY 235 ArgGluValValProAsnIleAspArgTyrIleAsnGlnThrPheLeuGlyLeuPro 254
 |||||
 Db 311 ATATACCTTCCTGACGAGTTCCTTGGAGCCTGGTACACCAAGTTCCTGGGTACCC 370
 QY 255 ValSerIleGlyAspArgCysLeuThrAsnTyrAlaThrAspLeuGly---LysThr 273
 |||||
 Db 371 TGTACTTTGGGGTACCGGACCTCACCACCGATGCTCAGCAGCTGATGCTACC 430
 QY 274 ValTyrGlnSerThrAlaLysCysIleThrAspValProAspLysMetSerThrTyrLeu 293
 |||||
 Db 431 AGTACACCTCCAGGTCCTGCTACTCAGACGACCCCTGCTCTCTGCGGTGCTG 490

```

QY 294 LysGlnGlnAsnArgTrrpAsnLysSerPhePheArgGluSerIleIleSerValLysLys 313
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 491 ACCACGACGACGCTGGCCAGTCTACTTCCTGAGTGGCTG----- 535
QY 314 IleMetAsnAsnPropheValAlaLeuTrp 323
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 536 -----TACAACGCGCTCTGG 550

RESULT 11
BJ041963          656 bp mRNA linear EST 07-DEC-2001
LOCUS
DEFINITION
  BJ041963 NIBB Mochii normalized Xenopus neurula library Xenopus
  laevis cDNA clone XL050c20 5', mRNA sequence.
ACCESSION
  BJ041963
VERSION
  BJ041963.1 GI:17417509
KEYWORDS
  EST.
SOURCE
  Xenopus laevis (African clawed frog)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
  Xenopodinae; Xenopus.
REFERENCE
  1 (bases 1 to 656)
AUTHORS
  Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i.T. and Kohara
  ,Y.
TITLE
  Expressed genes in X. laevis embryo
JOURNAL
  Unpublished
COMMENT
  Contact: Tadasu Shin-i
  Center For Genetic Resource Information
  National Institute of Genetics
  1111 Yata, Mishima, Shizuoka 411-8540, Japan
  Tel: 81-559-81-6855
  Fax: 81-559-81-6855
  Email: tshini@genes.nig.ac.jp.
FEATURES
    source
    1..656
    /organism="Xenopus laevis"
    /mol_type="mRNA"
    /db_xref="taxon:8355"
    /clone="XL050c20"
    /tissue_type="whole embryo"
    /dev_stage="stage 15"
    /clone_lib="NIBB Mochii normalized Xenopus neurula
    library"

BASE COUNT  159 a  169 c  142 g  184 t    2 others
ORIGIN
Alignment Scores:
Pred. No.:      1.83e-19      Length:      656
Score:          273.50      Matches:      74
Percent Similarity: 51.40%      Conservative: 36
Best Local Similarity: 34.58%      Mismatches: 89
Query Match:     12.75%      Indels:      16
DB:              12          Gaps:          6

US-09-469-200D-2 (1-417) x BJ041963 (1-656)

QY 206 TyrAspAsnAlaPheGlyValGluArgAlaGlnSerValThrGlyAsnIleLeuVal 225
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 33  TACTGGATGGGTTTAACTATGAGAGCATGCCAGTCTACTTTGGCTGTGCAATGC 92
QY 226 CysSerGlyProLeuSerValTyrArgGluValValProAsnIleAspArgTyr 245
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 93  ATTACGCGCGCTGGGGATGTACCGCAATTCCTCTCCACGAATTCATGAAGACTGG 152
QY 246 IleAsnGlnThrPheLeuGlyIleProValSerIleGlyAspAspArgCysLeuThrAsn 265
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 153 TACACCAAGAAATTTGGTTCCCGATGCGATTTGGGGAGACCGCTACCTACCAAT 212
QY 266 TyrAlaThrAspLeuGly---LysThrValTyrGlnSerThrAlaLysCysIleThrAsp 284
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 213 CGAGTTTGTGCTCTGGCTATGCCAACCAATACACGGCCAGATCCAAATGCTTACTGAA 272
QY 285 ValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrrpAsnLysSerPhePhe 304

```

```

      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 273 ACACCCACGAGTACCTACGCTGGCTCAACAGCAACGCGATGGAGCAAGTCTCTTC 332
QY 305 ArgGluSerIleIleSerValLysLysIleMetAsnAsnPropheValAlaLeuTrpThr 324
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 333 CGAGATGCTCTACAATTCATTGTGGTTCCATAACAT-----CACTTATGGATG 383
QY 325 IleLeuGluValSerMet-----PheMetMetLeuValTyrSerValVal 339
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 384 ACCTAGCAAGCTGTGATTACTGGATTCTTCTCTCTCATGCGCACTGTCATCCAG 443
QY 340 AspPheValGlyValAsnValArgGluPheAspTrpLeuArgValLeuAlaPheLeuVal 359
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 444 CTCTTCTACCGTGGAGATC-----TGG---AACATCTCTCTGTTCTTGTGCTG 488
QY 360 IleIlePheIleValAlaLeuCysArg---AsnIleHisTyrMetLeuLysHisProLeu 378
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 489 ACAGTACAACCTGTAGGCTTATCAAAATCTTCTCTTCTAGTGCCTCCGAGGGAACATA 548
QY 379 SerPheLeuLeuSerPropheTyrGlyValLeuHisLeuPheValLeuGlnProLeuLys 398
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 549 GTCATGCTTTCATGCTCTTCTACTCAGTGTATACATGTCACAGTTTACTACCTGCCAAA 608
QY 399 LeuTyrSerLeuPheThrIleArgAsnAlaAspTrrpGlyThr 412
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 609 ATGTTGCCATTGCC-ACCATCAACAGCGAGGTGGGGCACA 649

RESULT 12
BJ062498          536 bp mRNA linear EST 07-DEC-2001
LOCUS
DEFINITION
  BJ062498 NIBB Mochii normalized Xenopus tailbud library Xenopus
  laevis cDNA clone XL068c24 5', mRNA sequence.
ACCESSION
  BJ062498
VERSION
  BJ062498.1 GI:17424132
KEYWORDS
  EST.
SOURCE
  Xenopus laevis (African clawed frog)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
  Xenopodinae; Xenopus.
REFERENCE
  1 (bases 1 to 536)
AUTHORS
  Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i.T. and Kohara
  ,Y.
TITLE
  Expressed genes in X. laevis embryo
JOURNAL
  Unpublished
COMMENT
  Contact: Tadasu Shin-i
  Center For Genetic Resource Information
  National Institute of Genetics
  1111 Yata, Mishima, Shizuoka 411-8540, Japan
  Tel: 81-559-81-6856
  Fax: 81-559-81-6855
  Email: tshini@genes.nig.ac.jp.
FEATURES
    source
    1..536
    /organism="Xenopus laevis"
    /mol_type="mRNA"
    /db_xref="taxon:8355"
    /clone="XL068c24"
    /tissue_type="whole embryo"
    /dev_stage="stage 25"
    /clone_lib="NIBB Mochii normalized Xenopus tailbud
    library"

BASE COUNT  137 a  120 c  134 g  144 t    1 others
ORIGIN
Alignment Scores:
Pred. No.:      1.55e-19      Length:      536
Score:          273.00      Matches:      60
Percent Similarity: 58.00%      Conservative: 27
Best Local Similarity: 40.00%      Mismatches: 61
Query Match:     12.73%      Indels:      2
DB:              12          Gaps:          2

```


size 1.55 kb. Constructed by Life Technologies. Note: This is a Xenopus Gene Collection (XGC) library."

BASE COUNT 201 a 210 c 227 g 236 t 2 others

Alignment Scores:

Pred. No.:	4.5le-18	Length:	876
Score:	262.50	Matches:	71
Percent Similarity:	47.84%	Conservative:	40
Best Local Similarity:	30.60%	Mismatches:	86
Query Match:	12.24%	Indels:	35
DB:	13	Gaps:	7

US-09-469-200D-2 (1-417) x B0907404 (1-876)

```
QY 204 IleArgTyrAspAsnAlaPheGlyValGluArgAlaAlaGlnSerValThrGlyAsnIle 223
Db 4 CTGGCTTACTGGATGCGTTTAACTGGAGAGGGCTGCCAGTCTTACTTCGACTGGGTG 63
QY 224 LeuValCysSerGlyProLeuSerValTyrArgGluValValProAsnIleAsp 243
Db 64 TCCTGTATAGTGGACCTCTGGGAATGTACCGAACAACATTCCTCCAGGTGTTTGGAA 123
QY 244 ArgTyrIleAsnGlnPheLeuGlyIleProValSerIleGlyAspAspArgCysLeu 263
Db 124 GCCTGGTACAGACAGAAATTTTGGGAACCTATGTACTTTGGGAGATAGACATCTG 183
QY 264 ThrAsnTyrAlaThrAspLeuGly---LysThrValTyrGlnSerThrAlaLysCysIle 282
Db 184 ACAACCCAGTGTGCTAGCATGGGATATCGCACCAATATACCCACAATCCAGACATTC 243
QY 283 ThrAspValProAspLysMetSerThrTyrLeuLysGlnGlnAsnArgTrpAsnLysSer 302
Db 244 TCCGAACCTCCATCCCTGTATCTCCGGTGTGTGAACCAACCCGGTGGACCAAGTCC 303
QY 303 PhePheArgGluSerIleSerValLysLysIleMetAsnAsnProPheValAlaLeu 322
Db 304 TACTTCGAGAGTGGCTGTATATGCCAGTGTGGCGCACAGCAT-----CACATC 354
QY 323 TrpThrIleLeuGluValSerMetPheMetMetLeuValTyrSerValValAspPhe--- 341
Db 355 TGGATGACCATGAG-----TCTGTGGTGTCTTCATC 387
QY 342 -----PheValGlyAsnValArgGlu 348
Db 388 TTTCCTCTCTCATCTACCCACTGTTATCCGGCTCATCTATGCCGGCACCATATGGAAT 447
QY 349 PheAspTrpLeuArgValLeuAlaPheLeuValIleIlePheIleValAlaLeuCysArg 368
Db 448 GTTGTGTGGCTC-----CTCCTGTGCATCCAGATCATGTCTCTCTCAA 492
QY 369 AsnIleHis---TyrMetLeuLysHisProLeuSerPheLeuLeuSerProPheTyrGly 387
Db 493 TCCATCTATCGCTGCTGGCTCCGCGCAACTTCATTATGCTCCTGTGATGTCTCTACTCC 552
QY 388 ValLeuHisIlePheValLeuGlnProLeuLysLeuTyrSerLeuPheThrIleArgAsn 407
Db 553 ATGCTGTACATGACTGGGCTTCTGCCATCCAGTACTTTCCTTTGACCTTANACAAG 612
QY 408 AlaAspTrpGlyThr-----ArgLysLysLeuLeu 417
Db 613 ACCGGTTGGGGAACATCTGGCGCAAGAGATAGTA 648
```

Search completed: October 2, 2003, 16:10:26
Job time : 2829 secs